MATHEMATISCHE ZEITSCHRIFT

76. BAND

MATHEMATISCHE ZEITSCHRIFT

UNTER STÄNDIGER MITWIRKUNG VON

E. KAMKE
TÜBINGEN

R. NEVANLINNA
HELSINKI

F. K. SCHMIDT
HEIDELBERG

HERAUSGEGEBEN VON

H. WIELANDT
TÜBINGEN

WISSENSCHAFTLICHER BEIRAT

W. BLASCHKE A. E. INGHAM H. KNESER
W. MAGNUS O. PERRON G. PICKERT

76. BAND

BERLIN · GÖTTINGEN · HEIDELBERG
SPRINGER-VERLAG
1961

Unveränderter Nachdruck 1975

Springer-Verlag Berlin Heidelberg New York

Inhalt des 76. Bandes

Seite

AHMAD, Z. U., On the absolute Cesàro summability factors of infinite series 295

ANDRÉ, J., Über Parallelstrukturen. Teil I: Grundbegriffe 85

ANDRÉ, J., Über Parallelstrukturen. Teil II: Translationsstrukturen 155

ANDRÉ, J., Über Parallelstrukturen. Teil III: Zentrale t-Strukturen 240

ANDRÉ, J., Über Parallelstrukturen. Teil IV: T-Strukturen 311

AVANN, S. P., Distributive properties in semi-modular lattices 283

BABISTER, A. W., An inhomogeneous hypergeometric function 164

BEHRENS, E.-A., Distributiv darstellbare Ringe. II 367

BERZ, E., Lösung der linearen Differentialgleichung durch Transformation in einen Schiefkörper 179

BIERLEIN, D., Der Graph meßbarer Funktionen mit abstraktem Definitionsbereich . 468

CARTWRIGHT, M. L., and E. F. COLLINGWOOD, The radial limits of functions meromorphic in a circular disc 404

CIGLER, J., Hausdorffsche Dimensionen spezieller Punktmengen 22

EHLICH, H., Die positiven Lösungen der Gleichung $y^a - [y^a] = y^b - [y^b] = y^c - [y^c]$. 1

GROEMER, H., Lagerungs- und Überdeckungseigenschaften konvexer Bereiche mit gegebener Krümmung 217

HAKEN, W., Ein Verfahren zur Aufspaltung einer 3-Mannigfaltigkeit in irreduzible 3-Mannigfaltigkeiten 427

HEATON, R., A Note on Polynomial Cocycles 235

ISHIGURO, K., Zur Gibbsschen Erscheinung für das Kreisverfahren 288

KÜNNETH, H., Zu einem Satz von Hjelmslev über monotone Bogen 194

KUNZ, E., Differentialformen inseparabler algebraischer Funktionenkörper 56

LEICHTWEISS, K., Zur Riemannschen Geometrie in Grassmannschen Mannigfaltigkeiten 334

LOONSTRA, F., Darstellungen von Gruppen als homomorphes Bild 149

MACDONALD, I. D., On Certain Varieties of Groups 270

MARCUS, S., Les approximations diophantiennes et la catégorie de Baire 42

MILLER, J. B., A class of series involving integral kernels. I 257

NASTOLD, H.-J., Zum Dualitätssatz in inseparablen Funktionenkörpern der Dimension 1 75

RAYMOND, F., Local Cohomology Groups with Closed Supports 31

Seite

RIEGER, G. J., Über die Anzahl der Produktzerlegungen ganzer Zahlen 226

ROSENBERG, A., Blocks and Centres of Group Algebras 209

SABIDUSSI, G., Graph Derivatives 385

SANDERS, B. L., Concerning a Theorem of Ky Fan's 51

SCHMIDT, H., Über das Additionstheorem der zyklischen Funktionen 46

SCHMIDT, H., Zur Charakterisierung der zyklischen Funktionen durch eine Funktionalgleichung 402

SCHUBERT, H., Bestimmung der Primfaktorzerlegung von Verkettungen 116

TILLMANN, H. G., Distributionen als Randverteilungen analytischer Funktionen. II . 5

VOGEL, W., Lineare Programme und allgemeine Vertretersysteme 103

WAGNER, A., On collineation groups of projective spaces. I 411

EHLICH, H.
Math. Zeitschr. 76, 1–4 (1961)

Die positiven Lösungen der Gleichung
$y^a - [y^a] = y^b - [y^b] = y^c - [y^c]$

Von

HARTMUT EHLICH

1. Es bezeichne $[y]$ wie üblich die größte ganze Zahl kleiner oder gleich y und es sei $\{y\} = y - [y]$. Herr E. G. STRAUS hat kürzlich die Frage aufgeworfen, ob aus

$$y \geqq 0 \quad \text{und} \quad \{y^a\} = \{y^b\} = \{y^c\}$$

für paarweise verschiedene positive ganze Zahlen a, b, c stets $\{y^a\} = 0$ folgt[1] [1a]). Die Antwort liefert, wie in der vorliegenden Note bewiesen werden soll, der folgende

Satz 1. *Sind a, b, c paarweise verschiedene positive ganze Zahlen und ist d ihr größter gemeinsamer Teiler, so folgt aus*

$$\{y^a\} = \{y^b\} = \{y^c\} \tag{1}$$

für $0 \leqq y \leqq 1$

$$\{y\} = 0, \quad d.h. \quad y = 0 \quad oder \quad y = 1$$

und für $y > 1$

$$\{y^d\} = 0, \quad d.h. \quad y^d \text{ ist eine ganze Zahl.}$$

Daß für $y > 1$ der Exponent d nicht durch eine kleinere positive Zahl ersetzt werden kann, zeigt das Beispiel $y = 2^{\frac{1}{3}}$, $a = 15$, $b = 9$, $c = 6$; in diesem Fall ist $d = 3$,

$$\{(2^{\frac{1}{3}})^{15}\} = \{(2^{\frac{1}{3}})^{9}\} = \{(2^{\frac{1}{3}})^{6}\} = 0 \quad \text{und} \quad \{(2^{\frac{1}{3}})^{3}\} = 0,$$

aber für jeden Exponenten s mit $1 \leqq s < d$ wird $\{(2^{\frac{1}{3}})^s\} \neq 0$.

2. O.B.d.A. sei

$$a > b > c. \tag{2}$$

Ist $0 \leqq y \leqq 1$, so erhalten wir allein aus der Forderung $\{y^a\} = \{y^b\}$:

$$\{y^a\} - \{y^b\} = y^a - y^b - [y^a] + [y^b] = y^b(y^{a-b} - 1) = 0,$$

wegen (2) also $y = 0$ oder $y = 1$.

[1]) Report of the Institute in the theory of numbers. University of Colorado, Boulder, Colorado, June 21 – July 17, 1959, S. 336.

[1a]) *Zusatz bei der Korrektur:* Das gleiche Problem wird in ähnlicher Weise in FRED SUPNICK, H. J. COHEN und J. F. KESTON: On the powers of a real number reduced modulo one. Transactions A.M.S. **94**, 244–257 (1960) behandelt. Auf diese Arbeit wiesen mich freundlicher Weise die Herren LEWIS und LEKKERKERKER hin.

3. Im folgenden sei daher stets $y>1$. Genügt y der Gl. (1), so gibt es zwei nichtnegative ganze Zahlen m und n, nämlich $m=[y^a]-[y^c]$ und $n=[y^b]-[y^c]$, so daß y eine gemeinsame Nullstelle der beiden Polynome

$$\begin{cases} f(x)=x^a-x^c-m \\ g(x)=x^b-x^c-n \end{cases} \tag{3}$$

ist. Sind andererseits m und n nichtnegative ganze Zahlen und haben die beiden Polynome (3) eine gemeinsame Nullstelle $y>1$, so folgt

$$y^a-y^c-m=\{y^a\}-\{y^c\}+[y^a]-[y^c]-m=0$$
$$y^b-y^c-n=\{y^b\}-\{y^c\}+[y^b]-[y^c]-n=0.$$

Die beiden Ausdrücke $\{y^a\}-\{y^c\}$ und $\{y^b\}-\{y^c\}$ sind also ganz. Da aber für jede reelle Zahl $\tilde{y}$ nach Definition $0\leqq\{\tilde{y}\}<1$ gilt, folgt $\{y^a\}-\{y^c\}=0$ und $\{y^b\}-\{y^c\}=0$, also (1). Für $y>1$ ist daher Satz 1 gleichwertig mit

Satz 2. *Genügen die positiven ganzen Zahlen a, b, c der Ungleichung* (2), *ist d ist größter gemeinsamer Teiler, sind m und n nichtnegative ganze Zahlen und haben die beiden Polynome* (3) *eine gemeinsame Nullstelle $y>1$, so ist y^d eine ganze Zahl.*

4. Da die beiden Polynome f und g die gemeinsame Nullstelle $y>1$ besitzen, gilt

$$m>n\geqq 1. \tag{4}$$

Wäre nämlich $m\leqq n$, so folgte

$$0=f(y)-g(y)=y^a-y^b-(m-n)\geqq y^b(y^{a-b}-1),$$

nach (2) also $y\leqq 1$ im Widerspruch zur Voraussetzung $y>1$. Ebenso folgte aus $n=0$

$$g(y)=y^c(y^{b-c}-1)=0,$$

wegen (2) also wieder $y\leqq 1$.

Bei den Koeffizienten von f tritt ein Vorzeichenwechsel auf. $f(x)$ hat demnach höchstens eine positive Nullstelle. Wegen $f(0)=-m<0$ ist deshalb

$$f(x)<0 \quad \text{für} \quad 0\leqq x<y, \qquad f(y)=0, \qquad f(x)>0 \quad \text{für} \quad y<x. \tag{5}$$

Ist nun u irgendeine Nullstelle von $f(x)$, so ergibt sich einerseits

$$m=|u^a-u^c|\geqq|u|^a-|u|^c, \quad \text{also} \quad f(|u|)\leqq 0 \tag{6}$$

und andererseits

$$|u|^a=|m+u^c|\geqq m-|u|^c\geqq 2-|u|^c. \tag{7}$$

Wäre $|u|<1$, so folgte aus der letzten Ungleichung

$$1>|u|^a\geqq 2-|u|^c>1.$$

Zusammen mit (5) und (6) erhalten wir daher: *Ist u irgendeine Nullstelle von f, so ist $1\leqq|u|\leqq y$.*

5. Da die beiden Polynome $f(x) = x^a - x^c - m$ und $g(x) = x^b - x^c - n$ ganze rationale Koeffizienten und mindestens eine gemeinsame Nullstelle haben, gibt es ein Polynom $t(x) = t_r x^r + t_{r-1} x^{r-1} + \cdots + t_0$ mit ganz rationalen Koeffizienten t_ν ($\nu = 0, 1, \ldots, r$) höchstem Koeffizienten $t_r = 1$ und positiven Grad r, das f und g teilt[2]). Die Nullstellen von t sind dabei die gemeinsamen Nullstellen von f und g. Das Produkt der Beträge dieser Nullstellen ist gleich $|t_0|$. Nach **4** ist jeder dieser Beträge größer oder gleich 1 und für die Nullstelle y gilt nach Voraussetzung $y > 1$. Daher ist auch $|t_0| > 1$ und, da t_0 eine ganze Zahl ist, ist $|t_0| \geqq 2$.

Da $t(x)$ ein Teiler von $g(x)$ ist, teilt t_0 insbesondere n; wegen $n > 0$ folgt deshalb $n \geqq 2$ und somit nach (4) $m \geqq 3$. Statt (7) erhalten wir damit für die Beträge aller Nullstellen u von f die günstigere Abschätzung

$$|u|^a = |m + u^c| \geqq 3 - |u|^c, \tag{7a}$$

die zeigt, daß für alle Nullstellen u von f die Ungleichung $|u| > 1$ gültig ist.

6. Nach Definition ist $t(x)$ auch ein Teiler von

$$n f(x) - m g(x) = n x^a - m x^b + (m - n) x^c.$$

Wegen $m = t(x) \frac{f(x)}{t(x)} - x^c (x^{a-c} - 1)$ sind x^c und $t(x)$ teilerfremd[2]). $t(x)$ teilt daher das Polynom $h(x) = n x^{a-c} - m x^{b-c} + m - n$. $h(x)$ hat somit die Nullstelle $y > 1$. Außerdem ist $h(1) = 0$. Die Ableitung

$$h'(x) = x^{b-c-1} \left(n(a-c) x^{a-b} - m(b-c) \right)$$

besitzt genau eine positive Nullstelle; diese muß nach dem Satz von ROLLE zwischen 1 und y liegen. Da ferner $h(0) = m - n > 0$ ist, ergibt sich

$$h(x) \geqq 0 \quad \text{für} \quad 0 \leqq x \leqq 1 \quad \text{und} \quad x \geqq y, \quad h(x) < 0 \quad \text{für} \quad 1 < x < y. \tag{8}$$

Bezeichnet v irgendeine Nullstelle von $h(x)$; so ist

$$m - n = m v^{b-c} - n v^{a-c} = |m v^{b-c} - n v^{a-c}| \geqq m |v|^{b-c} - n |v|^{a-c}$$

oder

$$h(|v|) = n |v|^{a-c} - m |v|^{b-c} + m - n \geqq 0.$$

Nach (8) gilt also entweder $0 \leqq |v| \leqq 1$ oder $y \leqq |v|$.

7. Ist nun w eine Nullstelle von $t(x)$, so ist auch

$$f(w) = g(w) = h(w) = 0.$$

Für $|w|$ müssen daher die Ungleichungen

$$1 < |w| \leqq y \quad \text{und} \quad 0 \leqq |w| \leqq 1 \quad \text{oder} \quad y \leqq |w|$$

gelten, d.h. es muß $|w| = y$ oder $w = y e^{i\varphi}$ ($0 \leqq \varphi < 2\pi$) sein. Die möglichen Werte für φ wollen wir im folgenden bestimmen:

[2]) Vgl. etwa O. PERRON, Algebra I, 3. Aufl., 1951, §§ 36, 37, oder G. PICKERT, Einführung in die höhere Algebra, 1951, Kap. III, oder B. L. VAN DER WAERDEN, Moderne Algebra I, 4. Aufl., 1955, §§ 20–22.

Aus $f(y e^{i\varphi}) = 0$ folgt $y^a e^{i a \varphi} = y^c e^{i c \varphi} + m$ und durch Übergang zum Quadrat der Beträge

$$y^{2a} = y^{2c} + m^2 + 2m\, y^c \cos c\varphi,$$

während sich aus $f(y) = 0$

$$y^{2a} = y^{2c} + m^2 + 2m\, y^c$$

ergibt. Durch Subtraktion beider Gleichungen erhalten wir $\cos c\varphi = 1$, also $e^{i c \varphi} = 1$. Aus $f(y) = f(y e^{i\varphi}) = g(y) = g(y e^{i\varphi}) = 0$ erhalten wir dann weiter

$$\frac{1}{y^a}\left(f(y\, e^{i\varphi}) - f(y)\right) = e^{i a \varphi} - 1 = 0$$

$$\frac{1}{y^b}\left(g(y\, e^{i\varphi}) - g(y)\right) = e^{i b \varphi} - 1 = 0.$$

Ist $d = (a, b, c)$, so sind aber alle gemeinsamen Lösungen φ $(0 \leqq \varphi < 2\pi)$ der drei Gleichungen

$$e^{i a \varphi} = 1, \qquad e^{i b \varphi} = 1, \qquad e^{i c \varphi} = 1$$

die d Werte $\varphi_\nu = 2\pi \frac{\nu}{d}$, $\nu = 0, 1, \ldots, d-1$. $t(x)$ hat also genau die Nullstellen $w_\nu = y \exp\left(2\pi i \frac{\nu}{d}\right)$, $\nu = 0, 1, \ldots, d-1$, d.h. es ist

$$t(x) = x^r + t_{r-1} x^{r-1} + \cdots + t_0 = \prod_{\nu=0}^{d-1} (x - w_\nu) = x^d - y^d.$$

Da t ganze rationale Koeffizienten hat, ist — wie in Satz 2 behauptet — y eine ganze Zahl.

Tübingen, Mathematisches Institut der Universität

(Eingegangen am 11. Juli 1960)

TILLMANN, H. G.
Math. Zeitschr. 76, 5–21 (1961)

Distributionen als Randverteilungen analytischer Funktionen. II

Von

HEINZ GÜNTHER TILLMANN

Vor einigen Jahren zeigte G. KÖTHE [*1*, *2*], daß die stetigen linearen Funktionale auf dem Raum $\boldsymbol{H}(\overline{R})$ der auf der abgeschlossenen reellen Geraden $\overline{R}$ holomorphen und im Punkte ∞ verschwindenden Funktionen umkehrbar eindeutig durch die Paare $\tilde{u}=(\tilde{u}_+, \tilde{u}_-)$ von holomorphen Funktionen in der oberen bzw. unteren Halbebene dargestellt werden können. Dabei ist die Zuordnung gegeben durch

$$\tilde{u}(\zeta)=\frac{1}{2\pi i}\, u\left[\frac{1}{z-\zeta}\right], \quad \zeta\notin\overline{R}$$

$$u[f]=\int_{C^+}\tilde{u}_+(\zeta)\, f(\zeta)\, d\zeta-\int_{C^-}\tilde{u}_-(\zeta)\, f(\zeta)\, d\zeta=u_+[f]+u_-[f], \qquad f\in H(\overline{R}).$$

Hier sind C^+, C^- geschlossene Jordan-Kurven, welche alle singulären Stellen von f in der jeweiligen Halbebene einschließen.

Haben $\tilde{u}_+$, $\tilde{u}_-$ auf der reellen Geraden noch stetige Randwerte und verschwinden sie für $\zeta\to\infty$, so lassen sich die Integrationswege auf die reelle Gerade verlegen und das Funktional u kann mit der durch die Randwerte bestimmten Funktion $u(x)=\tilde{u}_+(x+i0)-\tilde{u}_-(x-i0)$ identifiziert werden. Da jedes Funktional u approximiert werden kann durch Funktionale, denen Funktionen mit stetigen Randwerten auf R entsprechen, bezeichnete G. KÖTHE die stetigen linearen Funktionale auf $\boldsymbol{H}(\overline{R})$ allgemein als Randverteilungen der ihnen zugeordneten holomorphen Funktionenpaare.

Nun können gewisse dieser Randverteilungen in natürlicher Weise mit Distributionen (im Sinne von L. SCHWARTZ [*3*]) identifiziert werden. So lassen sich etwa alle Distributionen auf R mit kompaktem Träger als Randverteilungen von holomorphen Funktionen auffassen [*5*].

Aber dieses Resultat läßt sich noch wesentlich verallgemeinern. In dieser Note wird gezeigt, daß alle Distributionen t, die Ableitungen endlicher Ordnung von Funktionen aus L^p, $1<p<\infty$, sind (d. h. $t\in(\mathcal{D}'_{L^p})$ im Sinne von SCHWARTZ [*4*], S. 56), als Randverteilungen analytischer Funktionen $\tilde{t}(\zeta)$ aufgefaßt werden können und diese $\tilde{t}$ werden funktionentheoretisch charakterisiert.

Das Prinzip dieser Zuordnung ist das folgende:

Die Distribution t (als stetiges lineares Funktional auf dem Raume $\mathcal{D}$ der beliebig oft differenzierbaren Funktionen mit kompaktem Träger) wird fortgesetzt zu einem stetigen linearen Funktional $\hat{t}$ auf einem $\mathcal{D}$ umfassenden

Raum L, in dem $\mathcal{D}$ wie auch der Raum $\boldsymbol{H}(\overline{R})$ dicht liegen und die Topologie von L gröber ist als die Topologien von $\mathcal{D}$ bzw. $\boldsymbol{H}(\overline{R})$. Dann induziert $\hat{t}$ auf $\boldsymbol{H}(\overline{R})$ ein stetiges lineares Funktional t_0 und t_0 und t bestimmen sich gegenseitig eindeutig. t_0 wiederum ist umkehrbar eindeutig durch die holomorphe Funktion $\tilde{t}(\zeta)=\frac{1}{2\pi i}t_0\left[\frac{1}{z-\zeta}\right]=\frac{1}{2\pi i}\hat{t}_x\left[\frac{1}{x-\zeta}\right]$ charakterisiert, wobei $\tilde{t}=(\tilde{t}_+,\tilde{t}_-)$ aus je einer Komponente $\tilde{t}_+$ in der oberen und $\tilde{t}_-$ in der unteren Halbebene besteht.

Die Indikatrix $\tilde{t}$ vermittelt eine Approximation der Distributionen t durch die analytischen Funktionen $t_\varepsilon(x)=\tilde{t}_+(x+i\varepsilon)-\tilde{t}_-(x-i\varepsilon)$ und in den Räumen L^p, $\mathcal{D}_{L^p}$, $(\mathcal{D}_{L^p})'$, $1<p<\infty$, bilden die Abbildungen T_ε: $t\to t_\varepsilon$, $\mathrm{Re}(\varepsilon)>0$ eine stetige Halbgruppe, insbesondere konvergiert t_ε gegen t im Sinne der starken Topologie dieser Räume.

Für Funktionen $g\in L^2$ liefert die Zerlegung

$$g=g_+-g_-,\qquad g_\pm(x)=g_\pm(x\pm i\,0)=\lim_{\varepsilon\to 0}\tilde{g}_\pm(x\pm i\,\varepsilon)$$

insbesondere eine orthogonale Zerlegung von $L^2=L^2_+\oplus L^2_-$, wobei also $L^2_+(L^2_-)$ aus Randverteilungen holomorpher Funktionen in der oberen (unteren) Halbebene besteht und übereinstimmt mit den Fourier-Transformierten von Funktionen mit Träger in der rechten (linken) Halbgeraden. Alle diese Resultate sind auf mehrere Variable übertragbar.

In einer folgenden Note werden diese Resultate benutzt, um auch alle gemäßigten Distributionen und Distributionen endlicher Ordnung als Randverteilungen analytischer Funktionen darzustellen, und zwar als Randverteilungen „langsam wachsender" analytischer Funktionen.

Dieser Zusammenhang zwischen Distributionen und analytischen Funktionen liefert einen neuen und im Hinblick auf zahlreiche Anwendungen, insbesondere in der mathematischen Physik, recht zweckmäßigen Zugang zur Distributionentheorie:

Man geht aus von einem Teilraum von $\boldsymbol{H}(\Omega-\overline{R})$ der aus Paaren „langsam wachsender" analytischer Funktionen besteht. Jedes solche Paar $\tilde{u}=(\tilde{u}_+,\tilde{u}_-)$ kann man dann als eine Distribution u ansehen und u ist Limes der Näherungen $u_\varepsilon(x)=\tilde{u}_+(x+i\varepsilon)-\tilde{u}(x-i\varepsilon)$. Insbesondere bestimmt u dann auch ein stetiges lineares Funktional auf den Raum $\mathcal{D}$ der beliebig oft differenzierbaren Funktionen: $u[\varphi]=\lim_{\varepsilon\to 0}u_\varepsilon[\varphi]=\lim_{\varepsilon\to 0}\int_R u_\varepsilon(x)\,\varphi(x)\,dx$, so daß sich der Zusammenhang mit dem Schwartzschen Aufbau der Distributionentheorie recht einfach herstellen läßt.

1. Die Räume $\boldsymbol{H}(\overline{R})$ und $\mathcal{D}_{L^p}(R)$

a) Es sei Ω die Riemannsche Zahlenkugel, R die reelle Gerade und $\overline{R}=R\cup\{\infty\}$.

$\boldsymbol{H}(\overline{R})$ ist der lineare Raum der in $\overline{R}$ lokalanalytischen Funktionen (die in ∞ verschwinden), in dem auf folgende Weise eine lokalkonvexe Topologie

eingeführt ist[1]). $\{U_n\}$ sei eine monotone Folge von Umgebungen von $\overline{R}$ in Ω mit

$$U_1 > \overline{U}_2 > U_2 > \cdots > \overline{R}, \qquad \overline{R} = \bigcap_1^\infty U_n, \quad \text{z.B.}$$

$$U_n = U_n^0 = \left\{z; |\operatorname{Im} z| < \frac{1}{n} \vee |z| > n \vee z = \infty\right\}. \tag{1}$$

$\boldsymbol{HB}(U_n)$ sei der B-Raum der in U_n analytischen, in ∞ verschwindenden und auf $\overline{U}_n$ stetigen Funktionen mit der Norm

$$\|f\|_{U_n} = \sup_{z \in U_n} |f(z)| = \max_{z \in \overline{U}_n} |f(z)|; \qquad \boldsymbol{U}(U_n, \varepsilon_n) = \{f; \|f\|_{U_n} < \varepsilon_n\}. \tag{2}$$

Die Topologie $\mathfrak{T}_H$ von $\boldsymbol{H}(\overline{R}) = \bigcup_1^\infty \boldsymbol{HB}(U_n)$ ist nun die feinste lokalkonvexe Topologie, die auf jedem $\boldsymbol{HB}(U_n)$ eine Topologie erzeugt, die gröber ist, als die dort durch (2) definierte Normtopologie. Sie wird erzeugt durch die Umgebungen

$$\boldsymbol{U}\{\varepsilon_n\} = \bigsqcap_1^\infty \boldsymbol{U}(U_n, \varepsilon_n) = \left\{f; f = \sum_1^m \alpha_j f_j, \sum |\alpha_j| \leqq 1, \|f_j\|_{U_j} < \varepsilon_j\right\}. \tag{3}$$

Diese Topologie ist unabhängig von der Folge $\{U_n\}$. Eine Menge von Funktionen aus $\boldsymbol{H}(\overline{R})$ ist genau dann beschränkt, wenn sie in einem $\boldsymbol{HB}(U_n)$ enthalten und dort beschränkt ist.

b) $\mathfrak{D}_{L^p} = \mathfrak{D}_{L^p}(R)$ ist der Raum aller komplexwertigen, beliebig oft differenzierbaren Funktionen $\varphi(x)$, deren sämtliche Ableitungen zum Raume $L^p = L^p(R)$ gehören:

$$\varphi \in \mathfrak{D}_{L^p} \leftrightarrow \int_{-\infty}^{+\infty} \left|\frac{d^n}{dx^n}\varphi(x)\right|^p dx < \infty, \qquad n = 0, 1, 2, \ldots. \tag{4}$$

Durch die Normen

$$\|\varphi\|_{p,k} = \left[\int \left|\frac{d^k\varphi}{dx^k}(x)\right|^p dx\right]^{1/p}, \qquad k = 0, 1, 2, \ldots \tag{5}$$

wird in $\mathfrak{D}_{L^p}$ eine lokalkonvexe Topologie $\mathfrak{T}_{(p)}$ definiert. Ein volles Umgebungssystem wird gebildet durch die

$$\boldsymbol{V}_{p,n,\varepsilon} = \{\varphi; \|\varphi\|_{p,k} < \varepsilon \text{ für } k = 0, 1, \ldots n\}. \tag{6}$$

c) Eine Funktion $f(z) \in \boldsymbol{H}(\overline{R})$ ist auf R beschränkt und geht für große z mindestens wie $M/|z|$ gegen Null, so daß die Einschränkung von $f(z)$ auf R zu $\mathfrak{D}_{L^p}$ gehört für jedes $p > 1$. Da mit f auch alle Ableitungen zu $\boldsymbol{H}(\overline{R})$ und damit zu L^p gehören, ist also $\boldsymbol{H}(\overline{R})$ in $\mathfrak{D}_{L^p}(R)$ enthalten. Durch Vergleich der Topologien erhalten wir:

Satz 1. *$\boldsymbol{H}(\overline{R})$ ist Teilraum von $\mathfrak{D}_{L^p}(R)$, $p > 1$. Die Topologie $\mathfrak{T}_H$ in $\boldsymbol{H}(\overline{R})$ ist feiner als die durch $\mathfrak{T}_{(p)}$ induzierte Topologie.*

[1]) Vgl. [1], [2], [5]. Dort wurde der Raum $\boldsymbol{H}(\overline{R})$ mit $\mathfrak{A}(\overline{R})$ bezeichnet.

Zum Beweise benutzt man für $f \in \boldsymbol{U}(U_n, \delta)$ mit $U_n = U_n^0$ (vgl. (1)) nach der Cauchyschen Koeffizientenabschätzung

$$(7) \qquad \begin{cases} |f^{(\nu)}(x)| \leqq \delta \cdot n^\nu \nu! & \text{für } |x| < n+1 \\ |f^{(\nu)}(x)| \leqq \dfrac{\delta \cdot n \cdot \nu!}{(|x|-n)^{\nu+1}} & \text{für } |x| \geqq n+1. \end{cases}$$

Integration der p-ten Potenzen liefert dann $\boldsymbol{U}(U_n, \delta_n) < \boldsymbol{V}_{p,m,\varepsilon}$ mit geeignetem $\delta_n = \delta_n(m, \varepsilon) > 0$. Da $\boldsymbol{V}_{p,m,\varepsilon}$ absolut konvex ist, gilt dann auch

$$\boldsymbol{U}\{\delta_n\} = \bigsqcap_1^\infty \boldsymbol{U}(U_n, \delta_n) < \boldsymbol{V}_{p,m,\varepsilon}.$$

Satz 2. *$\boldsymbol{H}(\bar{R})$ ist in $\mathscr{D}_{L^p}$ dicht: Zu jedem $f \in \mathscr{D}_{L^p}$ gibt es eine Folge $f_j \in \boldsymbol{H}(\bar{R})$, die bezüglich $\mathscr{T}_{(p)}$ gegen f konvergiert.*

Zum Beweise genügt es, zu zeigen, daß jede beliebig oft differenzierbare Funktion φ mit kompaktem Träger durch Funktionen aus $\boldsymbol{H}(\bar{R})$ approximierbar ist. Für ein solches φ ist die Fourier-Transformierte eine ganze analytische Funktion $\hat{\varphi}$ und es gilt die Umkehrformel:

$$(9) \qquad \varphi(x) = \int_{-\infty}^{+\infty} \hat{\varphi}(t) \exp(-2\pi i t x)\, dt.$$

Wir setzen

$$(10) \qquad \varphi_N(z) = \int_{-N}^{+N} \hat{\varphi}(t) \exp(-2\pi i t z)\, dt.$$

Durch Abschätzung der Fourier-Transformierten und ihrer Ableitungen zeigt man dann, daß $\|\varphi(x) - \varphi_N(x)\|_{p,k} \to 0$ geht für $N \to \infty$ für jedes natürliche k, d.h. $\varphi = \lim_{N\to\infty} \varphi_N$ bezüglich der Topologie $\mathscr{T}_{(p)}$. $\varphi_N(z)$ ist eine ganze analytische Funktion. Es gilt also in der Umgebung jedes endlichen Intervalls $(-A, +A)$ die Cauchysche Integralformel:

$$\varphi_N(z) = \frac{1}{2\pi i} \int_C \frac{\varphi_N(\zeta)}{\zeta - z}\, d\zeta.$$

Setzt man nun

$$(11) \qquad \varphi_{N,A}(z) = \frac{1}{2\pi i} \int_{-A-i}^{+A-i} \frac{\varphi_N(\zeta)}{\zeta - z}\, d\zeta - \frac{1}{2\pi i} \int_{-A+i}^{+A+i} \frac{\varphi_N(\zeta)}{\zeta - z}\, d\zeta,$$

so erhält man unter Ausnutzung der Eigenschaften der Fourier-Transformation die Abschätzung:

$$\|\varphi_N - \varphi_{N,A}\|_{p,k} \leqq M_k \cdot A^{-\frac{p-1}{p}},$$

so daß $\varphi_{N,A}$ für beliebiges $p > 1$ mit $A \to \infty$ gegen φ_N strebt im Sinne der Topologie $\mathscr{T}_{(p)}$. Da $\mathscr{D}_{L^p}$ ein F-Raum ist, kann man eine Diagonalfolge φ_{N_j, A_j} auswählen, die schon gegen φ konvergiert. (Im vorliegenden Fall genügt es, etwa $N_j = j$, $A_j = j^j$ zu setzen.) Aus Satz 1 und 2 folgt unmittelbar:

Korollar. *Ist t ein stetiges lineares Funktional in $\mathscr{D}_{L^p}$, so induziert es ein stetiges lineares Funktional t_0 in $\boldsymbol{H}(\bar{R})$ und ist durch t_0 eindeutig bestimmt.*

2. Die zu $H(\bar{R})$ und $\mathcal{D}_{L^p}(R)$ dualen Räume

Der zu $H(\bar{R})$ duale Raum ist nach G. KÖTHE [1] isomorph zum Raum $H(\mathbb{C}\bar{R})$ der in der oberen und unteren Halbebene holomorphen Funktionen. Die Isomorphie ist gegeben durch

$$(12)\quad \begin{cases} u \to \tilde{u}(\zeta) = \dfrac{1}{2\pi i}\, u\left[\dfrac{1}{z-\zeta}\right], \quad \operatorname{Im}\zeta \neq 0, \quad \tilde{u}(\zeta) = \begin{cases} \tilde{u}_+(\zeta), & \operatorname{Im}\zeta > 0 \\ \tilde{u}_-(\zeta), & \operatorname{Im}\zeta < 0 \end{cases} \\ u[f] = \int\limits_{C_+} \tilde{u}_+(\zeta) f(\zeta)\, d\zeta + \int\limits_{C_-} \tilde{u}_-(\zeta) f(\zeta)\, d\zeta, \qquad f \in H(\bar{R}). \end{cases}$$

$\tilde{u}$ nennt man die Indikatrix von u.

Dabei sind C_+ bzw. C_- geschlossene Kurven in oberer bzw. unterer Halbebene, die bei entgegengesetzter Orientierung gemeinsam ein unbeschränktes Gebiet beranden in dem f holomorph ist.

Für $\tilde{u}_+, \tilde{u}_- \in H(\bar{R})$ sind diese Funktionale völlig durch die Randwerte dieser Funktionen auf der reellen Achse bestimmt, weswegen man allgemein die stetigen Funktionale u auf $H(\bar{R})$ als Randverteilungen der Funktionen $\tilde{u}_\pm$ bezeichnet.

Die stetigen linearen Funktionale auf $\mathcal{D}_{L^p}$ sind Distributionen u (vgl. SCHWARTZ [4], S. 55 ff.), welche also zugleich Randverteilungen analytischer Funktionen $\tilde{u}(\zeta) \in H(\mathbb{C}\bar{R})$ sind. Aus der Tatsache, daß u als stetiges Funktional auf einer Nullumgebung von $\mathcal{D}_{L^p}$ beschränkt ist, ergibt sich eine Wachstumsbeschränkung der Funktionen $\tilde{u}(\zeta)$ bei Annäherung an die reelle Gerade:

Satz 3. *Ist $\tilde{u}(\zeta)$ eine in der oberen und unteren Halbebene holomorphe Funktion, die als Randverteilung eine auf $\mathcal{D}_{L^p}$ stetige Distribution besitzt, so wächst $\tilde{u}$ langsam bei Annäherung an die reelle Achse in folgendem Sinne: $\tilde{u}(\zeta)$ genügt einer Abschätzung*

$$(13)\quad |\tilde{u}(\zeta)| \leqq M \sup\left\{\eta^{+\frac{1}{p}-1}, \eta^{\frac{1}{p}-n}\right\}, \qquad \eta = |\operatorname{Im}\zeta|.$$

Beweis. u bildet eine Umgebung $V_{p,m,\varepsilon}$ in den Einheitskreis ab; also gilt

$$(14)\quad |\tilde{u}(\zeta)| = \frac{1}{2\pi}\left|u\left[\frac{1}{x-\zeta}\right]\right| \leqq \frac{1}{2\pi\varepsilon} \sup_{0\leq k\leq m} \left\|\frac{1}{x-\zeta}\right\|_{p,k},$$

also

$$|\tilde{u}(\zeta)| \leqq \frac{1}{2\pi\varepsilon} \sup_{0\leq k\leq m}\left[\int\limits_{-\infty}^{+\infty} \frac{(k!)^p\, dx}{|x-\zeta|^{(k+1)p}}\right]^{\frac{1}{p}}.$$

Nun ist

$$\int\limits_{-\infty}^{+\infty} \frac{dx}{|x-\zeta|^r} = 2\int\limits_0^\infty \frac{dx}{(x^2+\eta^2)^{r/2}} \leqq \frac{2}{\eta^{r-1}}\left(1+\frac{1}{r-1}\right),$$

also

$$(15)\quad |\tilde{u}(\zeta)| \leqq \frac{1}{2\pi\varepsilon} \sup_{0\leq k\leq m}\left\{k!\left(2+\frac{2}{kp+p-1}\right)^{\frac{1}{p}} \eta^{\frac{1}{p}-k-1}\right\}.$$

Das ist aber gerade die behauptete Abschätzung mit $\varkappa = m+1$. Nun gilt weiter, daß jede der Abschätzung (13) genügende „langsam wachsende" Funktion als Randverteilung eine Distribution besitzt.

Satz 4. *Sei $\tilde{u}(\zeta)$ eine in $\mathfrak{C}\overline{R}$ holomorphe Funktion, die der Abschätzung* (13) *genügt. Dann ist die Randverteilung von $\tilde{u}$ die m-te Ableitung einer stetigen Funktion $\varphi(x)$ die einer Abschätzung $|\varphi(x)| \leqq M'(1+|x|^{m-\delta})$ genügt und zu jedem x_0 ein Polynom höchstens $(m-1)$-ten Grades $p_{x_0}(x)$ existiert, so daß*

$$|\varphi(x) - p_{x_0}(x)| \leqq M'(1+|x-x_0|^{m-\delta}) \tag{16}$$

gilt[2]*). Dabei ist $m \geqq n$, $\delta = \frac{p-1}{p}$.*

Beweis. Wir betrachten die Komponente $\tilde{u}_-$ von $\tilde{u}$ in der unteren Halbebene. (Für $\tilde{u}_+$ verlaufen die Überlegungen ganz analog.)

Das durch $\tilde{u}_-(\zeta)$ bestimmte Funktional u_- wird gegeben durch

$$u_-[f] = -\int_{C_-} \tilde{u}_-(\zeta) f(\zeta)\, d\zeta, \qquad f \in \boldsymbol{H}(\overline{R}), \tag{17}$$

wobei C_- eine geschlossene Kurve ist, welche den Bereich der unteren Halbebene in dem f nicht definiert ist, im Uhrzeigersinne umläuft. Wir wählen C_- speziell so, daß sie aus der Strecke $s = [-a - i\varepsilon, b - i\varepsilon]$ und einem Kreisbogen K vom Radius $\frac{a+b}{2}$ besteht. Da für große ζ gilt $|f(\zeta)| \leqq \frac{M_f}{|\zeta|}$, folgt wegen (13)

$$\left| \int_K \tilde{u}_-(\zeta) f(\zeta)\, d\zeta \right| \leqq \int_K \frac{M \cdot M_f}{\eta^\delta \cdot |\zeta|} |d\zeta|, \qquad \delta = 1 - \frac{1}{p} = \frac{p-1}{p} > 0. \tag{18}$$

Das zweite Glied geht aber mit $a, b \to \infty$ gegen 0, so daß wir erhalten

$$u_-[f] = -\int_{-\infty - i\varepsilon}^{+\infty - \varepsilon i} \tilde{u}_-(\zeta) f(\zeta)\, d\zeta = -\int_{-\infty}^{+\infty} \tilde{u}_-(\zeta - i\varepsilon) \cdot f(\zeta - i\varepsilon)\, d\zeta. \tag{19}$$

Dieses Integral formen wir um durch partielle Integration und schätzen dazu zunächst das Integral von $\tilde{u}_-$ ab.

$$\tilde{u}_-^{(-1)}(\zeta; \xi_0) = \int_{\xi_0 - i}^{\zeta} \tilde{u}_-(\zeta)\, d\zeta; \qquad \zeta = \xi - i\eta. \tag{20}$$

Wählen wir den Integrationsweg $\xi_0 - i \to \xi_0 - i\varrho \to \xi - i\varrho \to \xi - i\eta = \zeta$ mit $\varrho = \operatorname{Max}\{1, |\xi - \xi_0|\}$, so erhalten wir für $\eta \leqq 1$

$$|\tilde{u}_-^{(-1)}(\zeta, \xi_0)| \leqq M_1 \{|\xi - \xi_0|^{1-\delta} + 1 + \eta^{1-\delta-(n-1)}\}. \tag{21}$$

Mit dem Integrationsweg $\xi_0 - i \to \xi - i \to \xi - i\eta = \zeta$ zeigt man dann durch Induktion, daß die iterierten Integrale $\tilde{u}_-^{(-k)}(\zeta; \xi_0) = \int_{\xi_0 - i}^{\zeta} \tilde{u}_-^{(-k+1)}(\zeta, \xi_0)\, d\zeta$ der

[2]) Man kann leicht zeigen, daß für jede stetige Funktion $\varphi(x)$, die (16) genügt, auch die Ableitung $u = \varphi^{(m)}$ (im Sinne der Distributionstheorie) Randverteilung einer gemäß (13) langsam wachsenden holomorphen Funktion $\tilde{u}(\zeta)$ in $\mathfrak{C}\overline{R}$ ist.

Abschätzung genügen:

$$|\tilde{u}_-^{(-k)}(\zeta, \xi_0)| \leqq M_k\{|\xi - \xi_0|^{k-\delta} + \sum_0^{k-1} |\xi - \xi_0|^{\nu} + \eta^{1-\delta-(n-k)}\}. \tag{22}$$

Für $k=n$ ergibt sich

$$\tilde{u}_-^{(-n)}|(\zeta; \xi_0)| \leqq M_n'\{|\xi - \xi_0|^{n-\delta} + 1\} \quad \text{in} \quad 0 < \eta \leqq 1 \tag{23}$$

und $\tilde{u}_-^{(-n)}(\zeta, \xi_0)$ ist auf die reelle Gerade noch stetig fortsetzbar. M_n' hängt dabei nicht von dem Anfangspunkt ξ_0 ab. Aus (19) erhalten wir nun durch partielle Integration

$$\left\{\begin{aligned} u_-[f] &= -\int_{-\infty}^{+\infty} \tilde{u}_-(\zeta - i\,\varepsilon)\, f(\zeta - i\,\varepsilon)\, d\zeta = (-1)^{n+1} \int_{-\infty}^{+\infty} \tilde{u}_-^{(-n)}(\zeta, \xi_0)\, f^{(n)}(\zeta)\, d\zeta \\ &= (-1)^{n+1} \int_{-\infty}^{+\infty} \tilde{u}_-^{(-n)}(x - i\,0, \xi_0)\, f^{(n)}(x)\, dx \end{aligned}\right. \tag{24}$$

und analog $u_+[f] = (-1)^n \int_{-\infty}^{+\infty} \tilde{u}_+^{(-n)}(x + i\,0, \xi_0)\, f^{(n)}(x)\, dx$, also

$$u[f] = u_+[f] + u_-[f] = \int_{-\infty}^{+\infty} (-1)^n \varphi_n(x, \xi_0)\, f^{(n)}(x)\, dx \tag{25}$$

mit

$$\varphi_n(x, \xi_0) = \tilde{u}_+^{(-n)}(x + i\,0; \xi_0) - \tilde{u}_-^{(-n)}(x - i\,0; \xi_0). \tag{26}$$

Da $\varphi_n(x; \xi_0) - \varphi_n(x, \xi_0')$ ein Polynom höchstens $(n-1)$-ten Grades ist, folgt die Behauptung des Satzes.

Setzt man $u_{\pm\varepsilon}(x) = \tilde{u}_\pm(x \pm i\,\varepsilon; \xi_0)$, so ist

$$\varphi_{n,\varepsilon}(x) = u_{+\varepsilon}^{(-n)}(x) - u_{-\varepsilon}^{(-n)}(x) = \tilde{u}_+^{(-n)}(x + i\,\varepsilon) - \tilde{u}_-^{(-n)}(x - i\,\varepsilon) \tag{27}$$

eine analytische Funktion auf R und

$$u_\varepsilon[f] = \int_{-\infty}^{+\infty} (-1)^n \varphi_{n,\varepsilon}(x)\, f^{(n)}(x)\, dx = \int_{-\infty}^{+\infty} \{\tilde{u}_+(x + i\,\varepsilon) - \tilde{u}_-(x - i\,\varepsilon)\} f(x)\, dx \tag{28}$$

konvergiert stark gegen $u[f]$:

$$\left\{\begin{aligned} |u_-[f] - u_{-,\varepsilon}[f]| &= |\int \{\tilde{u}_-^{(-n-1)}(x - i\,0) - \tilde{u}^{(-n-1)}(x - i\,\varepsilon)\} f^{(n+1)}(x)\, dx| \\ &\leqq \varepsilon \int |\tilde{u}_-^{(-n)}(x - i\,\varepsilon\,\vartheta_x)\, f^{(n+1)}(x)\, dx| \\ &\leqq \varepsilon \int M_n' (1 + |x|)^{n-\delta} |f^{(n+1)}(x)|\, dx. \end{aligned}\right. \tag{29}$$

Da sich $\tilde{u}_\pm^{(-n)}(x - i\eta) = \tilde{u}_-^{(-n)}(x - i\eta; 0)$ in dem ganzen Streifen $0 \leqq \eta \leqq \varepsilon$ durch $M_n'(|x|^{n-\delta} + 1)$, $f^{(n+1)}(x)$ für alle f aus einer beliebigen beschränkten Menge B in $H(\bar{R})$ durch $M_B \cdot \inf\{1, |x|^{-n-2}\}$ abschätzen läßt, geht $u_\pm[f] - u_{\pm,\varepsilon}[f]$ mit $\varepsilon \to 0$ auf jeder beschränkten Menge $B \subset H(\bar{R})$ gleichmäßig gegen Null. Aus (29) folgt auch unmittelbar, daß $u - u_\varepsilon$ auf jeder beschränkten Menge B aus dem Raum $\mathcal{D}$ (den beliebig oft differenzierbaren Funktionen mit kompaktem Träger) [vgl. [3], Chap. III, Th. IV, S. 70] gleichmäßig gegen Null geht. Wir haben damit bewiesen:

Satz 5. *Ist u Randverteilung einer gemäß (13) langsam wachsenden holomorphen Funktion $\tilde{u}(\zeta)$, so konvergieren die u_ε, die durch*

$$u_\varepsilon(x) = \tilde{u}_+(x+i\varepsilon) - \tilde{u}_-(x-i\varepsilon)$$

definiert sind, mit $\varepsilon \to 0$ gegen u im Sinne der starken Topologie des zu $\boldsymbol{H}(\overline{R})$ dualen Raumes und im Sinne der starken Topologie des Raumes $\mathcal{D}'$ der Distributionen auf R und u_ε läßt sich darstellen in der Form

$$(30)\quad \begin{cases} u_\varepsilon[f] = \int\limits_{-\infty}^{+\infty} \{\tilde{u}_+(x+i\varepsilon) - \tilde{u}_-(x-i\varepsilon)\} f(x)\,dx \\ \qquad = (-1)^n \int\limits_{-\infty}^{+\infty} \{\tilde{u}_+^{(-n)}(x+i\varepsilon) - \tilde{u}_-^{(-n)}(x-i\varepsilon)\} f^{(n)}(x)\,dx, \\ u[f] = \lim\limits_{\varepsilon\to 0} u_\varepsilon[f] \\ \qquad = (-1)^n \int \{\tilde{u}_+^{(-n)}(x+i0) - \tilde{u}_-^{(-n)}(x-i0)\} f^{(n)}(x)\,ax. \end{cases}$$

Wir wollen nun die Randverteilungen, die zu $(\mathcal{D}_{L^p})'$ gehören, genauer charakterisieren und beginnen dazu mit der Untersuchung derjenigen holomorphen Funktionen, welche eine Funktion aus L^q als Randverteilung haben. Sei

$$g \in L^q(R), \qquad \tilde{g}(\zeta) = \frac{1}{2\pi i}\int g(x)\,\frac{1}{x-\zeta}\,dx, \qquad 1<q<\infty.$$

Da g auf $\mathcal{D}_{L^p}$, $\frac{1}{p}+\frac{1}{q}=1$, ein stetiges lineares Funktional definiert, ist $\tilde{g}(\zeta)$ eine in $\mathfrak{C}\overline{R}$ holomorphe Funktion, die der Abschätzung (13) (mit $n=1$) genügt.

Für die ε-Näherung $g_\varepsilon(x) = \tilde{g}_+(x+i\varepsilon) - \tilde{g}_-(x-i\varepsilon)$ gilt

$$g_\varepsilon(x) = \frac{1}{2\pi i}\int g(y)\left[\frac{1}{y-x-i\varepsilon} - \frac{1}{y-x+i\varepsilon}\right]dy = \frac{\varepsilon}{\pi}\int g(y)\,\frac{1}{(x-y)^2+\varepsilon^2}\,dy,$$

also

$$g_\varepsilon = g * h_\varepsilon \quad \text{mit} \quad h_\varepsilon(t) = \frac{\varepsilon}{\pi}\,\frac{1}{t^2+\varepsilon^2}.$$

Es ist

$$\|h_\varepsilon\|_{L^1} = \frac{\varepsilon}{\pi}\int\limits_{-\infty}^{+\infty}\frac{dt}{t^2+\varepsilon^2} = \frac{2}{\pi}\int\limits_0^{\infty}\frac{dt}{t^2+1} = 1,$$

also $g_\varepsilon \in L^q$ und $\|g_\varepsilon\|_{L^q} \leqq \|g\|_{L^q}\cdot\|h_\varepsilon\|_{L^1} = \|g\|_{L^q}$.

Die Abbildungen $T_\varepsilon: g \to g_\varepsilon$ sind also stetige lineare Abbildungen von L^q in sich mit Norm $\|T_\varepsilon\| \leqq 1$. (Die T_ε, $\varepsilon>0$, bilden offenbar eine einparametrige Halbgruppe.)

Für $1<q<\infty$ ist $\boldsymbol{H}(\overline{R})$ in L^q dicht (da $\mathcal{D}_{L^q}$ in L^q dicht ist) und für $f \in \boldsymbol{H}(\overline{R})$ gilt:

$$f = f_+ - f_-, \qquad f_\pm \in \boldsymbol{H}(\overline{R} \cup \{\operatorname{Im}\zeta \gtrless 0\}),$$

$$f(x) - f_\varepsilon(x) = \{\tilde{f}_+(x) - \tilde{f}_+(x+i\varepsilon)\} - \{\tilde{f}_-(x) - \tilde{f}_-(x-i\varepsilon)\},$$

$$\tilde{f}_+(x) - \tilde{f}_+(x+i\varepsilon) = -i\varepsilon\,\tilde{f}_+'(x+i\varepsilon\vartheta_x), \qquad 0<\vartheta_x<1,$$

und

$$\|f(x)-f_\varepsilon(x)\|_{L^q}^q=\varepsilon^q\int|\tilde f'_+(x+i\,\varepsilon\,\vartheta_x)-\tilde f'_-(x-i\,\varepsilon\,\vartheta'_x)|^q\,dx.$$

Nun gilt für $\tilde f'(z)$ aber in einem Streifen $|\operatorname{Im} z|<\delta$ eine Abschätzung

$$|f'_\pm(z)|\leqq M(1+|z|^2)^{-1}\leqq M(1+x^2)^{-1},$$

also

$$\|f_+(x)-f_{+\varepsilon}(x)\|_{L^q}\leqq 2\,\varepsilon\,[\int M^q(1+x^2)^{-q}\,dx]^{1/q}=O(\varepsilon).$$

Da für $f_-(x)-f_{-,\varepsilon}(x)$ die gleiche Abschätzung gilt, folgt $\|f-f_\varepsilon\|_{L^q}\to 0$ für $\varepsilon\to 0$ für jedes $f\in\boldsymbol{H}(\bar R)$. Aus dem Satz von BANACH-STEINHAUS folgt also $\|g-g_\varepsilon\|_{L^q}\to 0$ mit $\varepsilon\to 0$ für jedes $g\in L^q$.

Damit haben wir

Satz 6. *Für $g\in L^q$, $0<q<1$, gilt*

$$(\alpha)\qquad |\tilde g(x\pm i\,y)|\leqq\frac{p}{\pi}\,\frac{\|g\|_{L^q}}{y^{1/p}}=:\frac{M}{y^{1/p}},\qquad \frac1p+\frac1q=1,$$

$$(\beta)\qquad \{g_\varepsilon(x)=\tilde g_+(x+i\,\varepsilon)-\tilde g_-(x-i\,\varepsilon);\ \varepsilon>0\}$$

ist eine beschränkte Menge in L^q, $(\|g_\varepsilon\|_{L^q}\leqq\|g\|_{L^q})$,

$$(\gamma)\qquad \|g-g_\varepsilon\|_{L^q}\to 0\quad \textit{mit}\quad \varepsilon\to 0.$$

Die Bedingungen (α) und (β) sind hinreichend dafür, daß $\tilde g$ eine Funktion aus L^q als Randverteilung hat.

Denn nach Satz 5 konvergieren die $g_\varepsilon(x)$ wegen (α) in $\mathscr{D}'$ stark (also auch schwach) gegen eine Distribution u.

Wegen (β) ist aber die Menge der g_ε in L^q relativ schwach kompakt, es gibt also eine Folge $\varepsilon_\nu\to 0$, derart, daß g_{ε_ν} in L^q schwach konvergiert gegen ein $g\in L^q$. Da aus schwacher Konvergenz in L^q aber schwache Konvergenz in $\mathscr{D}'$ folgt, muß $u=g$ sein.

Nun können wir eine genaue Charakterisierung der holomorphen Funktionen geben, welche Distributionen aus $(\mathscr{D}_{L^p})'$ als Randverteilungen haben.

Satz 7. *Jede Distribution $u\in(\mathscr{D}_{L^p})'$, $1<p<\infty$, ist Randverteilung einer holomorphen Funktion $\tilde u(\zeta)$, für die gilt:*

(a) $|\tilde u(\zeta)|\leqq M\sup\left(\eta^{-\frac{p-1}{p}},\eta^{-n+\frac1p}\right)$, $\qquad\eta=|\operatorname{Im}\zeta|$,

(b) $u_\varepsilon(x)=\tilde u_+(x+i\,\varepsilon)-\tilde u_-(x-i\,\varepsilon)\in(\mathscr{D}_{L^p})'$, $\{u_\varepsilon;\ \varepsilon>0\}$ *ist beschränkt in* $(\mathscr{D}_{L^p})'$

(c) $u_\varepsilon\to u$ *im Sinne der starken Topologie von* $(\mathscr{D}_{L^p})'$.

Beweis. u ist endliche Summe von Ableitungen von Funktionen aus L^q, $\frac1p+\frac1q=1$, also $u=\sum_0^m g_\nu^{(\nu)}$, $g_\nu\in L^q$

$$\tilde u(\zeta)=\sum_0^m\frac{d^\nu}{d\zeta^\nu}\,\tilde g_\nu(\zeta)$$

$$u_\varepsilon(x)=\sum_0^m\frac{d^\nu}{d\zeta^\nu}\,g_{\nu,\varepsilon}(x).$$

Nach Satz 6 gilt $g_{\nu,\varepsilon} \in L^q$, also $u_\varepsilon \in (\mathcal{D}_{L^p})'$ und $g_{\nu,\varepsilon} \to g_\nu$ in starker Topologie von L^q, also auch in der starken Topologie von $(\mathcal{D}_{L^p})' \supset L^q$. Da die Differentiation in $\mathcal{D}_{L^p}$ und damit in $(\mathcal{D}_{L^p})'$ stetig ist, folgt auch $u_\varepsilon \to u$ in der starken Topologie von $(\mathcal{D}_{L^p})'$.

Satz 8. *Jede holomorphe Funktion $\tilde{u}(\zeta)$, für die die Bedingungen* (a) *und* (b) *aus Satz 7 erfüllt sind, hat als Randverteilung eine Distribution $u \in (\mathcal{D}_{L^p})'$.*

Beweis. Wegen (a) konvergiert u_ε in $\mathcal{D}'$ schwach gegen eine Distribution u.

$\mathcal{D}_{L^p}$ ist ein reflexiver F-Raum. In $\mathcal{D}_{L^p}$ und $(\mathcal{D}_{L^p})'$ sind also alle beschränkten Mengen relativ schwach kompakt.

Wegen (b) kann man also eine Folge $\varepsilon_j \to 0$ finden, so daß u_{ε_j} in $(\mathcal{D}_{L^p})'$, also auch in $\mathcal{D}'$ gegen eine Element aus $(\mathcal{D}_{L^p})'$ konvergiert (schwach), also gilt $u \in (\mathcal{D}_{L^p})'$ und Satz 7 zeigt, daß $u_\varepsilon \to u$ sogar stark konvergiert.

Für den Fall $p = 2$ sind weitere Aussagen möglich.

Satz 9. *L^2 besitzt eine orthogonale Zerlegung $L^2 = L^2_+ \oplus L^2_-$, wobei L^2_+ aus Randverteilungen von holomorphen Funktionen in der oberen, L^2_- aus Randverteilungen von holomorphen Funktionen in der unteren Halbebene besteht. Insbesondere gilt für jedes $g \in L^2$*

$$g = g_+ - g_-, \qquad g_+ \in L^2_+, \qquad g_- \in L^2_-$$
$$\tilde{g}_+(x + i\varepsilon) \to g_+(x), \qquad \tilde{g}_-(x - i\varepsilon) \to g_-(x)$$

mit $\varepsilon \to 0$ im Sinne der Normtopologie von L^2. $L^2_\pm$ besteht aus allen Funktionen, die die Fourier-Transformierten von Funktionen mit Trägern in der positiven bzw. negativen Halbgeraden sind.

Beweis. Für g aus dem in L^2 dichten Teilraum $\boldsymbol{H}(\overline{R})$ gilt $g(z) = g_+(z) - g_-(z)$, wobei g_+ bzw. g_- auf der abgeschlossenen oberen bzw. unteren Halbebene holomorph sind. Es gilt

$$\int_{-\infty}^{+\infty} g_+(x)\,\overline{g_-(x)}\,dx = \int_{-\infty}^{+\infty} g_+(z)\,\overline{g_-(\bar{z})}\,dz = 0,$$

da der Integrand eine in der oberen Halbebene holomorphe Funktion ist, die für $|z| \to \infty$ wie $M/|z|^2$ gegen Null geht. Die Zerlegung $g = g_+ - g_-$ ist also eine orthogonale Zerlegung und $\|g\|^2 = \|g_+\|^2 + \|g_-\|^2$. Ist $\varphi = \lim g_j$ ein beliebiges Element aus L^2, $g_j \in \boldsymbol{H}(\overline{R})$, so gilt also $g_{j+} \to \varphi_+$, $g_{j-} \to \varphi_-$, wobei $\varphi_+, \varphi_- \in L^2$ und $\varphi = \varphi_+ - \varphi_-$ gilt. Da starke Konvergenz in L^2 aber starke Konvergenz in $\boldsymbol{H}(\overline{R})'$ zur Folge hat, konvergieren die $\tilde{g}_{j\pm}(\zeta)$ auf jeder kompakten Teilmenge von $\mathfrak{C}\overline{R}$ gleichmäßig gegen die Indikatrix von $\varphi_\pm$, d.h. φ_+, bzw. φ_- sind Randverteilungen von Funktionen aus der oberen bzw. unteren Halbebene.

Aus Satz 6 folgt dann $\tilde{\varphi}_+(x + i\varepsilon) \to \varphi_+(x)$ und $\tilde{\varphi}_-(x - i\varepsilon) \to \varphi_-$ im Sinne der L^2-Norm.

φ_+ geht bei der inversen Fourier-Transformation über in eine Funktion $\psi^+(y) = \int \varphi_+(z)\, e^{-2\pi i z y}\, dz$, die offenbar für $y < 0$ verschwindet, d.h. $\varphi_+ = \widehat{\psi^+}$ und analog $\varphi_- = \widehat{\psi^-}$, wobei $\psi^-(y)$ eine Funktion aus L^2 ist, die für $y > 0$ verschwindet.

Insbesondere ist also ψ^+ (bzw. ψ^-) Randverteilung einer holomorphen Funktion, die im Komplement von $[0, +\infty]$ (bzw. $[-\infty, 0]$) definiert ist und auf jeder Parallelen zur reellen Geraden quadratintegrabel ist.

Aus Satz 9 folgt eine analoge Zerlegung für die Räume $\mathcal{D}_{L^2}$ und $(\mathcal{D}_{L^2})'$:

Satz 10. *Es gilt die direkte topologische Zerlegung*

$$\mathcal{D}_{L^2} = \mathcal{D}_{L^2_+} \oplus \mathcal{D}_{L^2_-}, \qquad (\mathcal{D}_{L^2})' = (\mathcal{D}_{L^2_+})' \times (\mathcal{D}_{L^2_-})' = (\mathcal{D}_{L^2})'_- \times (\mathcal{D}_{L^2})'_+,$$

wobei $\mathcal{D}_{L^2_\pm}$ aus Funktionen besteht, deren sämtliche Ableitungen in $L^2_\pm$ liegen. $(\mathcal{D}_{L^2_\pm})' = (\mathcal{D}_{L^2})'_\mp$ besteht aus Randverteilungen von holomorphen Funktionen in der offenen Halbebene $\operatorname{Im}\zeta < 0$ bzw. $\operatorname{Im}\zeta > 0$. Für $g \in \mathcal{D}_{L^2}$ bzw. $u \in (\mathcal{D}_{L^2})'$ gehören die Komponenten $g_{\varepsilon\pm}$ zu $\mathcal{D}_{L^2_\pm}$ bzw. $u_{\varepsilon\pm}$ zu $(\mathcal{D}_{L^2})'_\pm$ und konvergieren im Sinne der Topologie von $\mathcal{D}_{L^2}$ bzw. der starken Topologie von $(\mathcal{D}_{L^2})'$ gegen $g_\pm$ bzw. $u_\pm$.

Beweis. Aus $\mathcal{D}_{L^2} < L^2$ folgt nach Satz 9 die Zerlegung $g = g_+ - g_-$ mit $g_+ \in L^2_+$, $g_- \in L^2_-$ und es gilt $g_{\varepsilon,+} = g_{+,\varepsilon} \to g_+$, $g_{-,\varepsilon} = g_{\varepsilon,-} \to g_-$. Da die Indikatrix von $g^{(\nu)}(x)$ gleich $\tilde{g}^{(\nu)}(\zeta)$ ist, folgt, daß $g_+ \in \mathcal{D}_{L^2}$ und $g_- \in \mathcal{D}_{L^2}$, sowie daß $g_{\varepsilon,\pm} \to g_\pm$ konvergiert im Sinne der Topologie von $\mathcal{D}_{L^2}$. Da g_+ und g_- in L^2 zueinander orthogonal sind, ist die Zerlegung eindeutig und die Projektionen $P_\pm: g \to g_\pm$ sind stetig, $\mathcal{D}_{L^2}$ ist also die topologische direkte Summe von $\mathcal{D}_{L^2_+}$ und $\mathcal{D}_{L^2_-}$.

Für ein Funktional $u \in (\mathcal{D}_{L^2})'$ gilt die Darstellung $u = \sum_0^m g_\nu^{(\nu)}$ mit $g_\nu \in L^2$, also die Zerlegung $u = \sum_0^m g_{\nu+}^{(\nu)} - \sum_0^m g_{\nu-}^{(\nu)} = u_+ + u_-$. Dabei gilt für $f = f_+ - f_- \in \mathcal{D}_{L^2}$

$$u[f] = u_+[f] + u_-[f] = \lim_{\varepsilon \to 0} \{u_{+,\varepsilon}[f] + u_{-,\varepsilon}[f]\} = u_+[-f_-] + u_-[f_+],$$

also

$$(\mathcal{D}_{L^2_+})' = (\mathcal{D}_{L^2})'_-; \qquad (\mathcal{D}_{L^2_-})' = (\mathcal{D}_{L^2})'_+.$$

Aus $g_{\nu,\varepsilon,\pm} \to g_{\nu,\pm}$ im Sinne der L^2-Norm für $g_\nu \in L^2$ folgt, da die Differentiation in $(\mathcal{D}_{L^2})'$ stetig ist, $u_{\varepsilon,\pm} = u_{\pm,\varepsilon} \to u_\pm$ im Sinne der Topologie von $(\mathcal{D}_{L^2})'$. Da die Topologie von $\mathcal{D}_{L^2}$ feiner ist als die Topologie des Raumes $\mathcal{E} = C^\infty$ aller beliebig oft differenzierbaren Funktionen, folgt für $g \in \mathcal{D}_{L^2}$ insbesondere die gleichmäßige Konvergenz von $g_{\varepsilon,\pm}^{(\nu)}$ gegen $g_\pm^{(\nu)}$ auf jedem kompakten Intervall.

Satz 11. *Ist $g \in L^p$, $1 \leqq p < \infty$, so gilt: $g_\varepsilon(x) = \tilde{g}(x + i\varepsilon) - \tilde{g}(x - i\varepsilon)$ konvergiert mit $\varepsilon \to 0$ gleichmäßig gegen $g(x)$ auf jedem kompakten Intervall, in dessen Umgebung $g(x)$ stetig ist.*

Beweis.

$$g_\varepsilon(x) = \frac{1}{2\pi i} \int_{-\infty}^{+\infty} g(y) \left\{ \frac{1}{y - x - i\varepsilon} - \frac{1}{y - x + i\varepsilon} \right\} dy = \frac{\varepsilon}{\pi} \int g(y) \frac{dy}{(x-y)^2 + \varepsilon^2}.$$

Wir zerlegen den Integrationsbereich in $I_1 = \{y: |x - y| \geqq \delta\}$ und $I_2 = \{y: |x - y| \leqq \delta\}$, $0 < \varepsilon < \delta < 1$. Es ist

$$\frac{\varepsilon}{\pi} \int_{I_1} |g(y)| \frac{dy}{(x-y)^2 + \varepsilon^2} \leqq \frac{2\varepsilon}{\pi \cdot \delta^2} \|g\|_{L^p}$$

nach Hölderscher Ungleichung. In I_2 zerlegen wir $g(y)=g_1(y)+ig_2(y)$ in Real- und Imaginärteil und können dann den Mittelwertsatz der Integralrechnung anwenden:

$$\frac{\varepsilon}{\pi}\int_{I_2} g(y)\,\frac{dy}{(x-y)^2+\varepsilon^2} = \{g_1(x_1)+i\,g_2(x_2)\}\,\frac{\varepsilon}{\pi}\int_{I_2}\frac{dy}{(x-y)^2+\varepsilon^2}$$

$$= \{g_1(x_1)+i\,g_2(x_2)\}\,\frac{1}{2\pi i}\left\{\oint\frac{dy}{y} - \int_{\delta-i\varepsilon}^{\delta+i\varepsilon}\frac{dy}{y} + \int_{-\delta-i\varepsilon}^{-\delta+i\varepsilon}\frac{dy}{y}\right\},\quad x_1, x_2\in I_2,$$

wobei das Kurvenintegral über das Rechteck

$$-\delta-i\varepsilon\to+\delta-i\varepsilon\to+\delta+i\varepsilon\to-\delta+i\varepsilon\to-\delta-i\varepsilon$$

zu erstrecken ist. Es ist

$$\int_{\pm\delta-i\varepsilon}^{\pm\delta+i\varepsilon}\left|\frac{dy}{y}\right| \leqq \frac{2\varepsilon}{\delta},$$

also mit einer Schranke M für $|g(y)|$ in $I_2=[x-\delta, x+\delta]$:

$$|g_\varepsilon(x)-g(x)| \leqq |g_1(x_1)-g_1(x)+i\,(g_2(x_2)-g_2(x))| + \frac{4\varepsilon M}{\pi\delta} + \frac{2\varepsilon}{\pi\delta^2}\|g\|_{L^p}.$$

Wir setzen nun etwa $\delta=\varepsilon^{\frac{1}{2}}$ und erhalten

$$|g_\varepsilon(x)-g(x)| \leqq |g_1(x_1)-g_1(x)+i\,(g_2(x_2)-g_2(x))| + \frac{4M}{\pi}\cdot\varepsilon^{\frac{1}{2}} + \frac{2\|g\|_{L^p}}{\pi}\varepsilon^{\frac{1}{2}}.$$

Sei nun $I=[a, b]$ ein kompaktes Intervall in dessen Umgebung $g(x)$ stetig, also gleichmäßig stetig ist. Da $x_1, x_2\in U_\delta$, $U_\delta=(a-\delta, b+\delta)$ ist wird also für alle $x\in I$ für jedes $\varepsilon^*>0$ $|g_\varepsilon(x)-g(x)|<\varepsilon^*$ wenn nur $\varepsilon=\delta^3$ hinreichend klein gewählt ist. Ist insbesondere $g(x)$ in $(-\infty, +\infty)$ beschränkt und gleichmäßig stetig, so ergibt sich gleichmäßige Konvergenz von $g_\varepsilon(x)$ gegen $g(x)$ auf der ganzen reellen Geraden und nicht nur auf kompakten Teilintervallen. Das gilt also insbesondere dann, wenn $g(x)$ eine stetige Funktion mit kompaktem Träger ist. (Der Satz wurde für den Fall stetiger Funktionen mit kompaktem Träger bewiesen in H. BREMERMANN-L. DURAND, [0]. Die Beweisführung konnte für den vorliegenden Fall von dort übernommen werden.)

3. Übertragung auf höhere Dimensionen

a) Es sei Ω^N der N-dimensionale Osgoodsche Raum (d.h. das N-fache kartesische Produkt der Zahlenkugel Ω), R^N der N-dimensionale reelle Zahlenraum und $\overline{R}^N=\overline{R}^N$ die abgeschlossene Hülle von R^N in Ω^N.

$\boldsymbol{H}(\overline{R}^N)$ sei der topologische Vektorraum der lokalanalytischen Funktionen in $\overline{R}^N$, die in allen unendlichen Punkten ihres Definitionsbereiches verschwinden (vgl. [5]; dort wurde dieser Raum mit $\mathfrak{A}(\overline{R}^N)$ bezeichnet). Die Topologie in $\boldsymbol{H}(\overline{R}^N)$ kann wieder durch (1) und (2) definiert werden, wenn man die Umgebungen U_n von $\overline{R}$ durch ihre N-fachen kartesischen Produkte U_n^N ersetzt, die dann eine Folge von Umgebungen von $\overline{R}^N$ mit $\bigcap_{n=1}^{\infty} U_n^N=\overline{R}^N$ bilden. Nach

[5] ist der zu $\boldsymbol{H}(\overline{R}^N)$ duale Raum isomorph dem Raum $\boldsymbol{H}((\Omega-\overline{R})^N)=\boldsymbol{H}((\mathbb{C}\overline{R})^N)$ der in $(\Omega-\overline{R})^N$ lokalanalytischen Funktionen. Die Isomorphie wird geliefert durch die Zuordnungen

$$\tilde{u}(\zeta)=(2\pi i)^{-N}\,u[(z_1-\zeta_1)^{-1}\dots(z_N-\zeta_N)^{-1}],\qquad \zeta=(\zeta_1,\dots,\zeta_N),\qquad \zeta_j\notin\overline{R},$$
$$u[f]=\int_{C_1}\cdots\int_{C_N} f(\zeta)\,\tilde{u}(\zeta)\,d\zeta,\qquad f\in\boldsymbol{H}(\overline{R}^N).$$

Dabei ist C_j die negativ orientierte Berandung einer Umgebung U_j von $\overline{R}$, derart, daß f in $U_1\times\cdots\times U_N$ holomorph ist. Bei dieser Zuordnung entspricht die starke Topologie der Topologie der gleichmäßigen Konvergenz auf allen kompakten Teilmengen von $(\Omega-\overline{R})^N$.

Die offene Menge $(\Omega-\overline{R})^N$ hat 2^N Komponenten G_σ, $\sigma=(\sigma_1,\dots,\sigma_N)$, $\sigma_j=\pm1$, $G_\sigma=\{\zeta:\sigma_j\operatorname{Im}\zeta_j>0\}$, so daß $\tilde{u}(\zeta)\in\boldsymbol{H}((\Omega-\overline{R})^N)$ aus 2^N Komponenten $\tilde{u}_\sigma$ besteht, wobei $\tilde{u}_\sigma$ in G_σ definiert ist. Es gilt also eine Zerlegung

$$u[f]=\sum_\sigma u_\sigma[f];\qquad u_\sigma[f]=\prod_{j=1}^N\sigma_j\int_{C_{\sigma_1}}\cdots\int_{C_{\sigma_N}} f(\zeta)\,\tilde{u}_\sigma(\zeta)\,d\zeta, \tag{33}$$

wobei C_{σ_j} eine geschlossene Kurve in $G_{\sigma_j}=\{\zeta_j:\sigma_j\operatorname{Im}\zeta_j>0\}$ ist, welche zur geschlossenen reellen Geraden $\overline{R}$ (mit üblicher Orientierung) homolog ist.

Haben die $\tilde{u}_\sigma$ stetige Randwerte auf R^N und verschwinden sie für $|\zeta|\to\infty$ (z.B. für $\tilde{u}_\sigma\in\boldsymbol{H}(\overline{R}^N)$, so kann man die Integrationswege auf den R^N verschieben und erhält:

$$\begin{cases} u[f]=\displaystyle\sum_\sigma\prod_{j=1}^N\sigma_j\int_{R^N}\cdots\int f(x)\,\tilde{u}_\sigma(x+i\,0\,\sigma)\,dx \\ \quad=\displaystyle\sum_\sigma\prod_{j=1}^N\sigma_j\int_{R^N}\cdots\int f(x)\lim_{\varepsilon\to0}\tilde{u}_\sigma(x+i\,\varepsilon\,\sigma)\,dx. \end{cases} \tag{34}$$

b) Der Raum $\mathscr{D}_{L^p}=\mathscr{D}_{L^p}(R^N)$ besteht (analog zum Fall $N=1$) aus allen beliebig oft differenzierbaren Funktionen $\varphi(x)=\varphi(x_1,\dots,x_N)$ deren sämtliche partiellen Ableitungen zu $L^p(R^N)$ gehören und die Topologie wird durch die Folge der Normen

$$\|\varphi\|_{p,k}=\Big[\int_{R^N}\cdots\int|D^k\varphi(x)|^p\,dx\Big]^{1/p};\qquad D^k=\Big(\frac{\partial}{\partial x_1}\Big)^{k_1}\cdots\Big(\frac{\partial}{\partial x_N}\Big)^{k_N}, \tag{35}$$

also durch die Nullumgebungen

$$V_{p,m,\varepsilon}=\{\varphi;\|\varphi\|_{p,k}<\varepsilon\ \text{für}\ |k|=k_1+\cdots+k_N\leqq m\} \tag{36}$$

erzeugt. ($\mathscr{D}_{L^p}$ ist also ein F-Raum.)

Ist $f\in\boldsymbol{H}(\overline{R}^N)$ und ist $f(z)$ in $(\overline{U_n^0})^N$ (vgl. (1)) definiert und dort $|f(z)|\leqq\delta$, so liefert die Cauchysche Integralformel die Abschätzung

$$|f^{(k)}(x)|=|D^kf(x)|\leqq\delta\cdot\prod_{j=1}^N g_j(x_j)\ \text{mit}\ g_j(t)=\begin{cases} n^{k_j} & \text{für}\ |t|\leqq n+1 \\ \dfrac{n}{(|t|-n)^{k_j+1}} & \text{für}\ |t|>n+1 \end{cases}, \tag{37}$$

woraus folgt, daß $\boldsymbol{H}(\overline{R^N})$ in $\mathcal{D}_{L^p}(R^N)$ enthalten ist und die Injektion stetig ist.

In $\mathcal{D}_{L^p}(R^N)$ ist der Raum $\mathcal{D}(R^N)$ der beliebig oft differenzierbaren Funktionen mit kompaktem Träger dicht und in $\mathcal{D}$ wiederum den Teilraum $\mathcal{D}_0$ aller Funktionen $\varphi(x)=\sum_{\nu=1}^{n}\prod_{j=1}^{N}\varphi_{j,\nu}(x_j)$, $\varphi_{j,\nu}\in\mathcal{D}(R)$. Nach Satz 2 können aber die Funktionen aus $\mathcal{D}_0$ im Sinne der Topologie von $\mathcal{D}_{L^p}$ durch Funktionen aus $\boldsymbol{H}(\overline{R^N})$ approximiert werden. Wir haben also die Übertragung von Satz 1 und 2:

Satz 12. *$\boldsymbol{H}(\overline{R^N})$ ist ein dichter Teilraum von $\mathcal{D}_{L^p}(R^N)$ und die Injektion $\boldsymbol{H}(\overline{R^N})\to\mathcal{D}_{L^p}(R^N)$ ist stetig für $1<p<\infty$.*

Korollar. *Jedes stetige lineare Funktional u auf $\mathcal{D}_{L^p}(R^N)$ induziert ein stetiges lineares Funktional auf $\boldsymbol{H}(\overline{R^N})$ und ist durch dieses eindeutig bestimmt. u kann also als Randverteilung holomorpher Funktionen angesehen werden.*

Da u auf einer Umgebung $V_{p,m,\varepsilon}$ beschränkt ist, genügt die Indikatrix $\tilde{u}(\zeta)$ einer Abschätzung

$$|\tilde{u}(\zeta)|=\frac{1}{(2\pi)^N}\left|u\left[\frac{1}{z_1-\zeta_1}\cdot\frac{1}{z_2-\zeta_2}\cdots\frac{1}{z_N-\zeta_N}\right]\right|\leqq M\cdot\sup_{0\leqq|k|\leqq m}\left\|\frac{1}{z_1-\zeta_1}\cdots\frac{1}{z_N-\zeta_N}\right\|_{p,k}$$

$$\leqq M\sup_{|k|\leqq m}k!\prod_{j=1}^{N}\left[\int_{-\infty}\frac{dx_j}{|x_j-\zeta_j|^{(k_j+1)p}}\right]^{1/p}\leqq M^*\sup_{|k|\leqq m}\prod_{j=1}^{N}\eta_j^{\frac{1}{p}-k_j-1}$$

also

$$|\tilde{u}(\zeta)|\leqq M^*\cdot\prod_{j=1}^{N}\left(\eta_j^{\frac{1}{p}-1}+\eta_j^{\frac{1}{p}-1-m}\right);\quad \eta_j=|\operatorname{Im}\zeta_j|,$$

in Verallgemeinerung von (13).

Hält man $(N-1)$ Variable fest, so hat man für die restliche Variable die Abschätzung (13). Indem man Satz 4 nacheinander auf die einzelnen Variablen $\zeta_1,\ldots,\zeta_N$ anwendet, erhält man also, daß nach endlich vielen Integrationen (etwa $m+1$ malige Integration nach jeder Variablen) $\tilde{u}_\sigma^{(-m_0\varepsilon)}(\zeta)$ stetige Randwerte auf R^N hat, und für das Funktional u die Darstellung gilt:

$$u=\sum u_\sigma$$

$$u_\sigma[f]=\Pi\sigma_j\int_{R^N+i\varepsilon\sigma}\cdots\int f(\zeta)\,\tilde{u}_\sigma(\zeta)\,d\zeta=\Pi\sigma_j\int_{R^N+i\varepsilon\sigma}\cdots\int(-1)^{m_0N}f^{(m_0\varepsilon)}(\zeta)\,\tilde{u}_\sigma^{(-m_0\varepsilon)}(\zeta)\,d\zeta$$

$$=\Pi\sigma_j\int_{R^N}\cdots\int(-1)^{m_0N}f^{(m_0\varepsilon)}(x)\,\tilde{u}_\sigma^{(-m_0\varepsilon)}(x+i\,0\,\sigma)\,dx$$

und die Näherungen $u_{\varepsilon,\sigma}=u_{\sigma,\varepsilon}$ mit $\tilde{u}_{\sigma,\varepsilon}(\zeta)=\tilde{u}_\sigma(\zeta+i\varepsilon\sigma)$, $\zeta\in G_\sigma$ konvergieren gleichmäßig gegen u_σ auf allen beschränkten Mengen von $\boldsymbol{H}(\overline{R^N})$ und auf allen beschränkten Mengen aus $\mathcal{D}(R^N)$. Damit haben wir die Übertragung von Satz 3, 4 und 5:

Satz 13. *Ist u ein stetiges lineares Funktional auf $\mathcal{D}_{L^p}(R^N)$, so genügt die Indikatrix $\tilde{u}(\zeta)$ von u einer Abschätzung*

$$|\tilde{u}(\zeta)|\leqq M\prod_{1}^{N}\left(\eta_j^{\frac{1}{p}-1}+\eta_j^{\frac{1}{p}-1-m}\right),\quad \eta_j=|\operatorname{Im}\zeta_j|. \tag{38}$$

Jede holomorphe Funktion $\tilde{u}(\zeta)$, *welche dieser Abschätzung genügt, hat nach einer endlichen Anzahl von Integrationen stetige Randwerte auf* R^N *und es gilt für das durch* $\tilde{u}$ *auf* $\boldsymbol{H}(\overline{R}^N)$ *definierte Funktional* $u=\Sigma u_\sigma$,

$$u_\sigma[f]=\Pi\sigma_j\int_{R^N}\cdots\int(-1)^{m_0 N} f^{(m_0 e)}(x)\,\tilde{u}_\sigma^{(-m_0 e)}(x+i\,0\,\sigma)\,dx,$$

also mit $\varphi(x)=\sum_\sigma \Pi\sigma_j\,\tilde{u}_\sigma^{(-m_0 e)}(x+i\,0\,\sigma)$:

$$\begin{aligned}u[f]&=\int_{R^N}\cdots\int(-1)^{m_0 N} f^{(m_0 e)}\left\{\sum_\sigma \Pi\sigma_j\,\tilde{u}_\sigma^{(-m_0 e)}(x+i\,0\,\sigma)\right\}dx\\&=\int_{R^N}\cdots\int(-1)^{m_0 e} f^{(m_0 e)}(x)\cdot\varphi(x)\,dx.\end{aligned}$$

u kann also auch als Distribution und zwar als eine Ableitung der stetigen Funktion $\varphi(x)$ *aufgefaßt werden. (Auf die genaue Charakterisierung der Funktionen* $\varphi(x)$ *verzichten wir, da wir sie im folgenden nicht benötigen.)*

Die Näherungen u_ε *mit* $\tilde{u}_{\varepsilon,\sigma}(\zeta)=\tilde{u}_\sigma(\zeta+i\,\varepsilon\sigma)$ *sind holomorphe Funktionen auf* R^N, *welche im Sinne der starken Topologie von* $(\boldsymbol{H}(\overline{R}^N))'$ *und der starken Topologie von* $(\mathscr{D}(R^N))'$ *gegen u konvergieren.*

Die genaue Charakterisierung der Indikatrixfunktionen $\tilde{u}(\zeta)$ für $u\in(\mathscr{D}_{L^p})'$ bietet gegenüber dem Fall $N=1$ keine neuen Probleme: Für $g\in L^q$ entsteht

$$g_\varepsilon(x)=\sum_\sigma(\Pi\sigma_j)\,\tilde{g}_\sigma(x+i\,\varepsilon\,\sigma)$$

aus $g(x)$ durch Faltung mit

$$h_\varepsilon^N(x)=\Pi\,h_\varepsilon(x_j)=\Pi\left(\frac{\varepsilon}{\pi}\,\frac{1}{x_j^2+\varepsilon^2}\right)\in L^1,$$

so daß $g_\varepsilon\in L^q$, $\|g_\varepsilon\|_{L^q}\leqq\|g\|_{L^q}$. Da wir für $g\in L^q\subset(\mathscr{D}_{L^p})'$, $\frac{1}{p}+\frac{1}{q}=1$, im Beweis von (38) $m=0$ setzen können, gilt $|\tilde{g}(\zeta)|\leqq M\prod_{j=1}^N\eta_j^{-\frac{1}{q}},\eta_j=|\operatorname{Im}\zeta_j|$. Aus dem Beweis von Satz 6 folgt schließlich für Funktionen $f(z)$ aus dem algebraischen Tensorprodukt $\boldsymbol{H}(\overline{R}_1)\otimes\cdots\otimes\boldsymbol{H}(\overline{R}_N)\subset\boldsymbol{H}(\overline{R}^N)$, daß in $\boldsymbol{H}(\overline{R}^N)$, also in $\mathscr{D}_{L^q}(R^N)$ und damit in $L^q(R^N)$ dicht liegt $\|f(x)-f_\varepsilon(x)\|_{L^q}\to 0$, woraus nach dem Satz von BANACH-STEINHAUS diese Aussage für alle $f\in L^q$ folgt:

Satz 14. *Für* $g\in L^q(R^N)$, $1<q<\infty$ *gilt:*

(α) $$|\tilde{g}(x\pm i\,y)|\leqq M\cdot\prod_1^N y_j^{-\frac{1}{q}}.$$

(β) $$\left\{g_\varepsilon(x)=\sum_\sigma\Pi\sigma_j\,\tilde{g}_\sigma(x+i\,\varepsilon\,\sigma)\right\}_{\varepsilon>0}$$

ist eine beschränkte Menge in L^q, $(\|g_\varepsilon\|_{L^q}\leqq\|g\|_{L^q})$, $g_\varepsilon(x)$ *ist holomorph in* R^N.

(γ) $$\|g-g_\varepsilon\|_{L^q}\to 0\quad mit\quad \varepsilon\to 0.$$

Die Bedingungen (α) *und* (β) *sind hinreichend dafür, daß* $\tilde{g}(\zeta)\in\boldsymbol{H}((\Omega-\overline{R})^N)$ *eine Funktion aus* L^q *als Randverteilung hat.*

Denn nach Satz 13 folgt aus (α), daß g_ε in $\mathscr{D}'$ stark, also auch schwach gegen eine Distribution u konvergiert. Wegen (β) ist aber die Menge der g_ε relativ schwach kompakt in L^q, d.h. $u=g\in L^q$ und $\tilde{g}$ ist die Indikatrix von g.

Daraus folgt dann insbesondere auch (γ).

Die Charakterisierung der Funktionen aus $\mathcal{D}_{L^q}(R^N)$ ist dann evident.

Da jedes stetige lineare Funktional $u \in (\mathcal{D}_{L^p})'$ als endliche Summe von Ableitungen von Funktionen aus $L^q(R^N)$ dargestellt werden kann, folgt auch leicht

Satz 15. *Jede Distribution* $u \in (\mathcal{D}_{L^p}(R^N))'$, $1 < p < \infty$, *ist Randverteilung einer holomorphen Funktion* $\tilde{u}(\zeta) \in \boldsymbol{H}((\Omega - \overline{R})^N)$, *für die gilt:*

(a) $|\tilde{u}(\zeta)| \leqq M \cdot \prod_1^N \left(\eta_j^{-\frac{1}{q}} + \eta_j^{-\frac{1}{q}-m}\right)$, $\eta_j = |\operatorname{Im} \zeta_j|$, $\frac{1}{p} + \frac{1}{q} = 1$.

(b) $\left\{u_\varepsilon(x) = \sum_\sigma \Pi\, \sigma_j \tilde{u}_\sigma(x + i\varepsilon\sigma)\right\}_{\varepsilon > 0}$ *ist beschränkt (also relativ schwach kompakt) in* $(\mathcal{D}_{L^p}(R^N))'$.

(c) $u_\varepsilon \to u$ *im Sinne der starken Topologie von* $(\mathcal{D}_{L^p})'$ *mit* $\varepsilon \to 0$.

Jede holomorphe Funktion $\tilde{u}(\zeta)$, *für die* (a) *und* (b) *gilt, ist die Indikatrix einer Distribution aus* $(\mathcal{D}_{L^p}(R^N))'$.

Die Schlußweise ist genau analog zu den Beweisen von Satz 7 und 8, da auch $\mathcal{D}_{L^p}(R^N)$ ein reflexiver F-Raum ist.

Für $f \in \boldsymbol{H}(\overline{R^N})$ ist die Zerlegung $f = \Sigma f_\sigma$, $f_\sigma(x) = \Pi \sigma_j \tilde{f}_\sigma(x + i0\sigma)$ eine Zerlegung in paarweise orthogonale Komponenten, da für $\sigma_j \neq \sigma_j'$ gilt

$$\int_{R_j} f_\sigma(x)\, \overline{g_{\sigma'}(x)}\, dx_j = \int_{R_j} f_\sigma(z)\, \overline{g_{\sigma'}(\bar{z})}\, dz_j = 0,$$

da der Integrand in der Halbebene $\sigma_j \cdot \operatorname{Im} \zeta_j \geqq 0$ holomorph ist und im Unendlichen eine zweifache Nullstelle hat. Da $\boldsymbol{H}(\overline{R^N})$ in $L^2(R^N)$ dicht ist, gilt auch für alle $g \in L^2$ die orthogonale Zerlegung

$$g = \Sigma g_\sigma, \quad g_\sigma(x) = \lim_{\varepsilon \to 0} g_{\varepsilon,\sigma}(x) = \lim_{\varepsilon \to 0} \tilde{g}_\sigma(x + i\varepsilon\sigma)$$

(Konvergenz im Sinne der L^2-Norm).

Satz 16. $L^2(R^N)$ *besitzt die orthogonale Zerlegung* $L^2 = \sum_\sigma \oplus L^2_\sigma$, *wobei* L^2_σ *aus allen Funktionen aus* L^2 *besteht, welche Randverteilungen von holomorphen Funktionen* $\tilde{u}_\sigma$ *in* G_σ *sind. Für jedes* $g \in L^2$ *gilt also* $g = \Sigma g_\sigma$, *wobei* $g_\sigma(x) = \lim_{\varepsilon \to 0} \tilde{g}_\sigma(x + i\varepsilon\sigma)$ *(Limes im Sinne der Normtopologie) aus* L^2_σ *ist.* L^2_σ *besteht aus den Fourier-Transformierten von Funktionen* $\psi^\sigma \in L^2$, *deren Träger in dem Bereich* $\{x; \sigma_j x_j \geqq 0,\ j = 1, \ldots, N\}$ *liegt.*

Die Aussage über die Fourier-Transformationen läßt sich auf die eindimensionale Fourier-Transformation zurückführen und folgt dann aus Satz 9.

Aus der Zerlegung von L^2 folgt wieder unmittelbar

Satz 17. *Es gilt die direkte topologische Zerlegung*

$$\mathcal{D}_{L^2}(R^N) = \sum_\sigma \mathcal{D}_{L^2_\sigma}(R^N); \qquad (\mathcal{D}_{L^2}(R^N))' = \prod_\sigma (\mathcal{D}_{L^2_\sigma}(R^N))' = \prod_{-\sigma} (\mathcal{D}_{L^2}(R^N))'_{-\sigma},$$

wobei $\mathcal{D}_{L^2_\sigma}$ aus den Funktionen besteht, deren sämtliche Ableitungen in L^2_σ liegen, $(\mathcal{D}_{L^2})'_\sigma$ besteht aus denjenigen Funktionalen u auf $\mathcal{D}_{L^2}$, deren Indikatrix $\tilde{u}(\zeta)$ nur in G_σ von Null verschieden ist. Der zu $\mathcal{D}_{L^2_\sigma}$ duale Raum ist der Raum $(\mathcal{D}_{L^2})'_{-\sigma}$. Die Zerlegungen $g=\Sigma g_\sigma \in \mathcal{D}_{L^2}$ bzw. $u=\{u_\sigma\}\in(\mathcal{D}_{L^2})'$ sind gegeben durch

$$g_\sigma(x)=\lim_{\varepsilon\to 0}\tilde{g}_\sigma(x+i\varepsilon\sigma), \qquad u_\sigma(x)=\lim_{\varepsilon\to 0}\tilde{u}_\sigma(x+i\varepsilon\sigma),$$

wobei die Konvergenz im Sinne der starken Topologie von $\mathcal{D}_{L^2}$ bzw. $(\mathcal{D}_{L^2})'$ erfolgt.

Für $g\in\mathcal{D}_{L^p}$ folgt aus der Konvergenz der Näherungen g_ε gegen g im Sinne der Topologie von $\mathcal{D}_{L^p}$ insbesondere die gleichmäßige Konvergenz von $g_\varepsilon(x)$ gegen $g(x)$ auf allen kompakten Intervallen.

Allgemeiner gilt:

Satz 18. *Ist $g\in L^p(R^N)$, $1\leqq p<\infty$, so gilt: $g_\varepsilon(x)=\sum_\sigma \Pi\sigma_j\tilde{g}_\sigma(x+i\varepsilon\sigma)$ konvergiert gleichmäßig gegen $g(x)$ auf jedem kompakten Bereich, in dessen Umgebung $g(x)$ stetig ist.*

Beweis. Es genügt die gleichmäßige Konvergenz zu beweisen auf allen kompakten Quadern in deren Umgebung $g(x)$ stetig ist. Es sei

$$(x|y_\nu)=(x_1,\ldots,x_{\nu-1}\,y_\nu,x_{\nu+1},\ldots,x_N), \qquad h_\varepsilon(t)=\frac{\varepsilon}{\pi}\frac{1}{t^2+\varepsilon^2}$$

$$g_{(\varepsilon),1}(x)=\int g\left(x|y_1\right)h_\varepsilon(x_1-y_1)\,dy_1$$

$$g_{(\varepsilon),\nu+1}(x)=\int g_{(\varepsilon),\nu}\left(x|y_{\nu+1}\right)h_\varepsilon(x_{\nu+1}-y_{\nu+1})\,dy_{\nu+1}, \qquad \nu=1,2,\ldots,N-1,$$

also

$$g_\varepsilon(x)=g_{(\varepsilon),N}(x).$$

Dann ist mit $g(x)=g_{(\varepsilon),0}(x)$

$$g_\varepsilon(x)-g(x)=\sum_1^N\left(g_{(\varepsilon),\nu}-g_{(\varepsilon)\nu-1}\right).$$

Es genügt also, die gleichmäßige Konvergenz von $g_{(\varepsilon),\nu}(x)-g_{(\varepsilon),\nu-1}(x)$ gegen Null zu zeigen, d.h. die Behauptung folgt unmittelbar aus Satz 11 und seinem Beweis, da die $g_{(\varepsilon),\nu}$ in den Quadern Q, in denen $g(x)$ stetig ist, gleichmäßig beschränkt und gleichstetig sind.

Literatur

[0] Bremermann, H. J., and L. Durand III.: On analytic continuation, multiplication and Fourier-transforms of Schwartz-distributions. Forschungsbericht, Institut for Advanced Study, Princeton N. J., Mai 1959. Erscheint in Trans. Amer. Math. Soc.

[1] Köthe, G.: Dualität in der Funktionentheorie. J. reine angew. Math. **191**, 30–49 (1953).

[2] Köthe, G.: Die Randverteilungen analytischer Funktionen. Math. Z. **58**, 13–33 (1952).

[3] Schwartz, L.: Théorie des distributions, I. Paris 1950.

[4] Schwartz, L.: Théorie des distributions, II. Paris 1951.

[5] Tillmann, H. G.: Randverteilungen analytischer Funktionen und Distributionen. Math. Z. **59**, 61–83 (1953).

Heidelberg, Mathematisches Institut der Universität

(Eingegangen am 16. Juli 1960)

CIGLER, J.
Math. Zeitschr. 76, 22–30 (1961)

Hausdorffsche Dimensionen spezieller Punktmengen

Von

JOHANN CIGLER

1. In den letzten Jahren erschienen mehrere Arbeiten, in denen die Hausdorffsche Dimension von Mengen reeller Zahlen untersucht wurde, deren g-adische Entwicklung gewissen Einschränkungen unterworfen ist (vgl. VOLKMANN [8] und die dort zitierte Literatur). Speziell interessierte man sich für Mengen von Zahlen, bei denen für die einzelnen Ziffern feste Frequenzen vorgeschrieben sind. Die Ergebnisse, die man dabei erzielte, legten die Vermutung nahe, daß eine enge Beziehung zwischen der Hausdorffschen Dimension einerseits und der Entropie der durch die Ziffernverteilung in der g-adischen Darstellung bestimmten "stationären ergodischen Quelle" andererseits besteht. Dieser Zusammenhang mit informationstheoretischen Fragestellungen wurde zuerst von KINNEY [5] aufgedeckt. Er lieferte nicht nur eine neue Einsicht in den Begriff der Hausdorffschen Dimension, sondern führte gleichzeitig auch zu einer Vereinfachung der Beweise verschiedener Sätze. Die Kinneysche Methode wurde dann in einer kürzlich erschienen Arbeit von BILLINGSLEY [1] weiter verfolgt und bedeutend ausgebaut. BILLINGSLEY definiert einen Begriff der Hausdorffschen Dimension in beliebigen Wahrscheinlichkeitsräumen, die gewissen Einschränkungen genügen, und gelangt dann zu Resultaten, die eine große Anzahl von bekannten Sätzen im klassischen Fall verallgemeinern und erweitern.

In der vorliegenden Arbeit werden einige Ergebnisse von VOLKMANN [8] verallgemeinert und einige spezielle Probleme untersucht. Es zeigt sich, daß ein interessanter Zusammenhang mit ϑ-adischen Entwicklungen besteht. Für diese Fragestellungen erweist sich eine Kombination der Methoden von KINNEY und VOLKMANN als vorteilhaft, da damit die Beweise besonders einfach und durchsichtig werden.

2. Die Hausdorffsche Dimension $\dim E$ einer Menge E ist folgendermaßen definiert: Sei $\eta>0$ eine reelle Zahl und $V_\eta=\{o_\nu\}$ eine η-Überdeckung von E mit offenen Mengen o_ν, d.h. es sei $E\subseteqq\bigcup_1^\infty o_\nu$ und das Lebesguesche Maß jeder dieser Mengen o_ν kleiner als η ($|o_\nu|<\eta$). Dann heißt der Grenzwert

$$\{M\}^\alpha=\lim_{\eta\to 0}\inf_{V_\eta}\sum_{\nu=1}^{\infty}|o_\nu|^\alpha \qquad (0<\alpha\leqq 1),$$

wo das Infimum über alle η-Überdeckungen von E zu erstrecken ist, das α-dimensionale Hausdorffsche Maß von E. Es gibt nun eine eindeutig bestimmte Zahl δ, so daß $\{M\}^\alpha=\infty$ für $\alpha<\delta$ und $\{M\}^\alpha=0$ für $\alpha>\delta$ gilt. Diese Zahl $\delta=\dim E$ heißt die Hausdorffsche Dimension von E.

Nun noch einige Konventionen über die zu verwendenden Bezeichnungsweisen: Unter g verstehen wir im folgenden stets eine ganze Zahl $g \geqq 2$. Wir betrachten alle reellen Zahlen $x \in [0, 1]$ in ihrer g-adischen Entwicklung:

$$x = \sum_{\nu=1}^{\infty} \frac{x_\nu}{g^\nu} = 0, \; x_1 x_2 x_3 \ldots .$$

Unter einem Grundintervall n-ter Stufe I_n verstehen wir ein Intervall der Gestalt

$$\frac{a_1}{g} + \frac{a_2}{g^2} + \cdots + \frac{a_n}{g^n} \leqq x < \frac{a_1}{g} + \frac{a_2}{g^2} + \cdots + \frac{a_n + 1}{g^n}.$$

Mit $J_n(x)$ bezeichnen wir das Grundintervall n-ter Stufe, das die Zahl x enthält. Ferner sei $A_{a_1 a_2 \ldots a_k}(x, n)$ die Anzahl der x_ν in der Entwicklung von x mit $1 \leqq \nu \leqq n - k + 1$, für die $x_\nu = a_1$, $x_{\nu+1} = a_2$, …, $x_{\nu+k-1} = a_k$ gilt. Existiert der Grenzwert

$$\lim_{n \to \infty} \frac{1}{n} A_{a_1 a_2 \ldots a_k}(x, n) = \zeta,$$

so nennen wir ζ die Frequenz der Ziffernsequenz $[a_1 a_2 \ldots a_k]$ in der Entwicklung von x.

3. Wir benötigen folgendes kombinatorische

LEMMA[1]. *Es bezeichne* $A = A(a_0, a_1, \ldots, a_{g-1};\; b_{00}, b_{01}, \ldots, b_{g-1,g-1})$ *die Anzahl aller* n*-tupel* $x_1 x_2 \ldots x_n$ *von ganzen Zahlen* x_ν *mit* $0 \leqq x_\nu \leqq g - 1$, *in denen die Ziffer* i *genau* a_i*-mal auftritt* $(i = 0, 1, \ldots, g-1)$ *und außerdem die Ziffernkombination* ij *genau* b_{ij}*-mal auftritt* (ij *durchläuft alle möglichen Werte*). *Dann gilt für* A *die folgende Abschätzung nach oben:*

$$A(a_0, \ldots, a_{g-1};\; b_{00}, \ldots, b_{g-1\,g-1}) \leqq (g-1)! \prod_{i=0}^{g-1} \prod_{j=0}^{g-1} \frac{a_i!}{b_{ij}!}.$$

BEWEIS. Sei $x_1 x_2 \ldots x_n$ ein festes n-tupel. Wir wollen zwei Ziffern x_ν und x_μ äquivalent nennen, wenn $x_\nu = x_\mu$ und außerdem $x_{\nu+1} = x_{\mu+1}$ gilt, wenn also die beiden Ziffernpaare $x_\nu x_{\nu+1}$ und $x_\mu x_{\mu+1}$ übereinstimmen. Um das zu erreichen, hängen wir an x_ν die Zahl $x_{\nu+1}$ als Index an und schreiben das n-tupel in der Gestalt: $x_{1 x_2} x_{2 x_3} \ldots x_{n x_{n+1}}$. An die letzte Ziffer kann jede der Zahlen $0, 1, \ldots, g-1$ als Index angehängt werden. Wir können nun die Ziffern unseres n-tupels folgendermaßen ordnen: Das erste Element sei $x_{1 x_2}$. Dann kommen alle $x_{\nu x_{\nu+1}}$, für die $x_\nu = x_1$ ist, und zwar in derselben Reihenfolge, wie sie im gegebenen n-tupel auftreten. Als nächstes Element wählen wir das erste $x_{\nu x_{\nu+1}}$, das noch nicht verwendet wurde, und führen damit dasselbe Verfahren durch, usw. Wir bekommen auf diese Art ein n-tupel aus a_{i_1} Ziffern i_1, a_{i_2} Ziffern i_2, …, a_{i_g} Ziffern i_g. Wieviel derartige n-tupel gibt es? Ein festes System von a_i Ziffern i kann offenbar in $\prod_{j=0}^{g-1} \frac{a_i!}{b_{ij}!}$ verschiedene nichtäquivalente Anordnungen übergeführt werden. Macht man

[1]) Vgl. auch BILLINGSLEY [1].

das bei jedem i und kombiniert die Anordnungen unabhängig voneinander, so gibt das $\prod_{i=0}^{g-1}\prod_{j=0}^{g-1}\frac{a_i!}{b_{ij}!}$ verschiedene Möglichkeiten. Nun können noch die $i_1, i_2, \ldots, i_g$ untereinander permutiert werden. Dies ist auf $g!$ Arten möglich. Da man jedem n-tupel $x_1 x_2 \ldots x_n$ genau g geordnete n-tupel zuordnen kann (wegen der Mehrdeutigkeit der Zuordnung bei x_n) und verschiedene n-tupel auch verschiedene geordnete n-tupel liefern, folgt daraus unmittelbar die Behauptung unseres Lemmas.

4. Wir beweisen zunächst zwei Sätze[2]), die im wesentlichen von KINNEY [5] stammen.

SATZ 1. *Sei μ ein normiertes Maß auf $[0, 1]$ und E die Menge aller $x \in [0, 1]$, für die bei festem $\alpha \geqq 0$ die Ungleichung*

$$\mu\left(I_n(x)\right) > g^{-n(\alpha+\varepsilon)} \tag{1}$$

bei beliebigem $\varepsilon > 0$ für unendlich viele n erfüllt ist. Dann gilt $\dim E \leqq \alpha$.

BEWEIS. Sei $C_n = \{I_j\}$ eine g^{-n}-Überdeckung von E durch Grundintervalle I_j und $\varepsilon > 0$ fest. Für jedes $x \in E$ wählen wir dasjenige Grundintervall $I_m(x)$, für das

$$\mu\left(I_m(x)\right) > |I_m(x)|^{\alpha+\varepsilon} = g^{-m(\alpha+\varepsilon)}$$

für das kleinste $m \geqq n$ erfüllt ist. Ein solches m gibt es nach Voraussetzung. Klarerweise sind die I_j fremd. Wegen $E \subseteqq \bigcup_j I_j$ gilt somit

$$1 = \mu([0, 1]) \geqq \mu\left(\bigcup_j I_j\right) = \sum_j \mu(I_j) > \sum_j |I_j|^{\alpha+\varepsilon}.$$

Läßt man n gegen Unendlich streben, so gilt also $\{E\}^{\alpha+\varepsilon} \leqq 1$, d. h. $\dim E \leqq \alpha + \varepsilon$. Da $\varepsilon > 0$ vollkommen beliebig war, ist also sogar $\dim E \leqq \alpha$, wzbw.

SATZ 2. *Es sei μ ein normiertes Maß auf $[0, 1]$ und E die Menge aller $x \in [0, 1]$, für die bei festem $\alpha \geqq 0$ und beliebigem $\varepsilon > 0$*

$$\mu\left(I_n(x)\right) < g^{-n(\alpha-\varepsilon)} \tag{2}$$

für alle hinreichend großen n gilt. Ist überdies $\mu(E) = 1$, dann gilt $\dim E \geqq \alpha$.

BEWEIS. Verläuft ganz analog wie bei KINNEY [5].

5. Die beiden folgenden Sätze bilden eine Verallgemeinerung von Resultaten von VOLKMANN [8].

SATZ 3. *Sei E die Menge aller Zahlen $x \in [0, 1]$, für die bei beliebigem $\varepsilon > 0$*

$$p_i p_{ij} - \varepsilon \leqq \frac{A_{ij}(x, n)}{n} \leqq p_i p_{ij} + \varepsilon \qquad (0 \leqq i, j \leqq g-1) \tag{3}$$

für unendlich viele n erfüllt ist, wobei $p_i \geqq 0$, $p_{ij} \geqq 0$, $\sum_j p_{ij} = p_i$, $\sum p_i = 1$ gilt. Dann ist

$$\dim E \leqq \alpha = -\frac{1}{\log g}\sum_{i=0}^{g-1}\sum_{j=0}^{g-1} p_i p_{ij} \log p_{ij}.$$

[2]) Satz 1 und 2 sind eine für unsere Zwecke vorteilhafte Umformulierung von Theorem I der Arbeit von KINNEY, wobei wir die Beschränkung auf Markowsche Ketten, die sich als überflüssig erweist, fallen lassen.

BEWEIS. Bezeichnet man mit $H_n(\varepsilon)$ die Menge aller x, die (3) erfüllen, so gilt $E \subseteqq \bigcup_{n=N}^{\infty} H_n(\varepsilon)$ für jedes N.

Die Anzahl $B_n(\varepsilon)$ der Grundintervalle n-ter Stufe, aus denen sich $H_n(\varepsilon)$ zusammensetzt, genügt der Ungleichung

$$B_n(\varepsilon) \leqq (2\varepsilon n + 1)^{g^2} \cdot A(p_0 n, \ldots, p_{g-1} n;\ p_0 p_{00} n, \ldots, p_{g-1} p_{g-1\,g-1} n),$$

wobei A die im Lemma eingeführte Größe bedeutet.

Nach unserem Lemma gilt also

$$B_n(\varepsilon) \leqq (2\varepsilon n + 1)^{g^2} (g-1)! \prod_{i=0}^{g-1} \prod_{j=0}^{g-1} \frac{(p_i n)!}{(p_i p_{ij} n)!}.$$

Wir wollen dabei annehmen, daß alle $p_i n$ und $p_i p_{ij} n$ ganz sind. (Wären sie nicht ganz, so müßte man sie durch ganze Zahlen derselben Größenordnung ersetzen; das ändert nichts an der nachfolgenden Abschätzung.) Aus der Stirlingschen Formel folgt dann

$$B_n(\varepsilon) \leqq e^{-n \sum_i \sum_j p_i p_{ij} \log p_{ij} + \varepsilon_1 n} = g^{n(\alpha+\varepsilon_2)}$$

mit

$$\lim_{\varepsilon \to 0} \varepsilon_2 = 0.$$

Für $\beta > \alpha$ ist also

$$\{E\}^{\beta} \leqq \lim_{N \to \infty} \sum_{n=N}^{\infty} \frac{g^{n(\alpha+\varepsilon_2)}}{g^{n\beta}} = \lim_{N \to \infty} \sum_{n=N}^{\infty} g^{n(\alpha-\beta+\varepsilon_2)} = 0$$

bei hinreichend kleinem $\varepsilon > 0$. Daraus ergibt sich, daß $\dim E \leqq \beta$ ist und da β eine beliebige Zahl $> \alpha$ war, auch die behauptete Relation $\dim E \leqq \alpha$.

Für den nächsten Satz benötigen wir noch einige Definitionen. Wir betrachten für jedes $x \in [0, 1]$ die Folge der Punkte

$$\mathfrak{y}_n(x) = \left(\frac{A_{00}(x, n)}{n-1}, \frac{A_{01}(x, n)}{n-1}, \ldots, \frac{A_{g-1,g-1}(x, n)}{n-1} \right) \qquad (n = 1, 2, 3, \ldots)$$

in einem g^2-dimensionalem euklidischen Raum.

Mit $V(x)$ bezeichnen wir die Menge aller Häufungspunkte der Folge $\{\mathfrak{y}_n(x)\}$. Man zeigt dann ganz analog wie bei VOLKMANN [8], daß die Menge $V(x)$ für jedes x ein Kontinuum, d.h. eine zusammenhängende abgeschlossene Menge, ist. Unter H verstehen wir die Menge aller Punkte $\mathfrak{x} = (x_{00}, x_{01}, \ldots, x_{g-1,g-1})$ des R^{g^2}, welche die Relationen $0 \leqq x_{ij} \leqq 1$ und $\sum_i \sum_j x_{ij} = 1$ erfüllen.

Ferner sei $d(\mathfrak{x}) = -\frac{1}{\log g} \sum_i \sum_j x_i x_{ij} \log x_{ij}$.

Dann gilt der

SATZ 4. *Sei S eine beliebige Menge mit $S \subseteqq H$ und $C(S)$ die Menge aller x mit $V(x) \subseteqq S$. Dann ist*

$$\dim C(S) \leqq \max_{\mathfrak{x} \in S} d(\mathfrak{x}).$$

BEWEIS. Sei $\mathfrak{z} \in S$ ein Punkt, in dem die Funktion $d(\mathfrak{x})$ ihr Maximum annimmt, und U die Menge aller Punkte $\mathfrak{x} \in H$ mit $d(\mathfrak{x}) \leqq d(\mathfrak{z}) + \varepsilon$. Dann

gilt $S \subsetneqq U$ und $C(S) \subsetneqq C(U)$. Die Anzahl $K(n)$ der Grundintervalle n-ter Stufe, in denen Elemente aus $C(U)$ liegen, ist

$$K(n) \leqq \sum_{i} \sum_{j} \prod_{i} \prod_{j} \frac{y_i!}{y_{ij}!} (g-1)!,$$

wobei die Summe über alle ganzen Zahlen y_i, y_{ij} zu erstrecken ist, für die $0 \leqq y_i \leqq n$, $0 \leqq y_{ij} \leqq n$ und $y_i = \sum_{j=0}^{g-1} y_{ij}$ gilt und für die außerdem die Punkte $\left(\frac{y_{00}}{n}, \frac{y_{01}}{n}, \ldots, \frac{y_{g-1,g-1}}{n}\right)$ in U liegen. Daraus folgt

$$K(n) \leqq n^{g^2} \max_{y_{ij}} (g-1)! \prod_{i} \prod_{j} \frac{y_i!}{y_{ij}!} \leqq g^{n(d(\mathfrak{g})+\varepsilon+\varepsilon_1)}.$$

Derselbe Schluß wie bei Satz 3 führt auch hier zum erwünschten Resultat.

6. Wir wollen nun die vorangegangenen Sätze auf einige spezielle Probleme anwenden.

SATZ 5. *Sei E die Menge aller Zahlen $x \in [0, 1]$, für die für alle $i\,j$ gilt*

$$\lim_{n \to \infty} \frac{A_{ij}(x, n)}{n} = p_i p_{ij}.$$

Dabei seien die p_i, p_{ij} derart, daß die Markowsche Kette mit den p_{ij} als Übergangswahrscheinlichkeiten ergodisch ist (vgl. A. I. KHINCHIN [4]). *Dann ist*

$$\dim E = \alpha = -\frac{\sum_{i} \sum_{j} p_i p_{ij} \log p_{ij}}{\log g}.$$

BEWEIS. Die Abschätzung nach oben ergibt sich unmittelbar aus Satz 3. Für die Abschätzung nach unten verwenden wir Satz 2. Wir betrachten die Menge $E^* \subset E$ aller Zahlen $x \in [0, 1]$, $x = \sum_{1}^{\infty} \frac{x_\nu}{g^\nu}$, in denen die x_ν eine Markowsche Kette mit den angegebenen Übergangswahrscheinlichkeiten bilden. Sei μ das Maß, das durch diese Markowsche Kette auf $[0, 1]$ induziert wird. Da die Kette ergodisch ist, gilt $\mu(E^*) = 1$ und $\mu\left(I_n(x)\right) = g^{-n\alpha+o(n)}$ für jedes $x \in E^*$, also speziell $\mu\left(I_n(x)\right) < g^{-n(\alpha-\varepsilon)}$ (vgl. KINNEY [5]). Damit sind alle Voraussetzungen zu Satz 2 erfüllt, und es gilt daher $\dim E \geqq \alpha$, also insgesamt $\dim E = \alpha$.

BEMERKUNG 1. Satz 1 liefert uns nur $\dim E^* = \alpha$. Es gibt jedoch nicht abzählbar viele Zahlen $x \in [0, 1]$, für die zwar die Voraussetzungen zu Satz 5 erfüllt sind, bei denen die x_ν jedoch keine Markowsche Kette bilden[3]). Man kann also mit Hilfe von Satz 1 keine Aussage über $\dim E$ machen.

BEMERKUNG 2. Satz 5 ist eine Verallgemeinerung eines Resultates von EGGLESTON [3], der den Fall betrachtet, wo die Frequenzen für jede einzelne Ziffer vorgegeben sind.

[3]) Es existieren nicht einmal immer die Frequenzen für Ziffernsequenzen $a_1 a_2 \ldots a_k$ mit $k \geqq 3$.

7. Bekanntlich kann man jede reelle Zahl $x \in [0, 1]$ eindeutig (bis auf eine abzählbare Menge) nach einer beliebigen Basis $\vartheta > 1$ entwickeln: $x = \sum_1^\infty \frac{x_\nu}{\vartheta^\nu}$ mit ganzen Zahlen x_ν, $0 \leqq x_\nu \leqq [\vartheta]$. Jede Folge $\{x_\nu\}$, die in der ϑ-adischen Entwicklung einer Zahl x auftreten kann, wollen wir ϑ-zulässig nennen.

Sei nun $g \geqq 1 + [\vartheta]$ eine ganze Zahl. Wir betrachten dann die Abbildung φ, die der Zahl $y = \sum_1^\infty \frac{x_\nu}{\vartheta^\nu}$ die Zahl $x = \sum_1^\infty \frac{x_\nu}{g^\nu}$ zuordnet. Dabei geht das Intervall $[0, 1]$ in eine Menge $E \subseteqq [0, 1]$ über. Auf $[0, 1]$ existiert ein bezüglich der Transformation $Tx = \vartheta x - [\vartheta x]$ invariantes ergodisches Maß ν, das äquivalent zum Lebesgueschen Maß ist (vgl. z.B. J. CIGLER [2]).

Durch die Festsetzung $\mu\left(I_n(x)\right) = \nu\left(\varphi^{-1}(I_n(x))\right)$, wenn $I_n(x)$ als Bildmenge auftritt, und $\mu\left(I_n(x)\right) = 0$ sonst wird auf $[0, 1]$ ein Maß μ definiert. Für alle $x \in E$ gilt $\mu\left(I_n(x)\right) \leqq c\vartheta^{-n}$ mit einer von x unabhängigen Konstanten c. Das folgt daraus, weil für ν die Relation $\nu(A) \leqq c|A|$ gilt, wobei A eine beliebige meßbare Menge bedeutet und $|A|$ das Lebesguesche Maß der Menge A, und weil die Länge eines ϑ-adischen Grundintervalles n-ter Stufe stets $\leqq 1/\vartheta^n$ ist. Für jedes $x \in E$ gilt also insbesondere $\mu\left(I_n(x)\right) \leqq g^{-n\left(\frac{\log \vartheta}{\log g} - \varepsilon\right)}$ für alle n.

Aus Satz 2 folgt somit der

SATZ 6. *Sei E die Menge aller Zahlen $x \in [0, 1]$, $x = \sum_1^\infty \frac{x_\nu}{g^\nu}$, wo die $\{x_\nu\}$ alle ϑ-zulässigen Folgen durchlaufen für ein $\vartheta > 1$ mit $[\vartheta] \leqq g - 1$. Dann gilt*

$$\dim E \geqq \frac{\log \vartheta}{\log g}.$$

Für die Abschätzung nach oben brauchen wir die Relation (1) für unendlich viele n mit $\alpha = \frac{\log \vartheta}{\log g}$. Diese ist sicher dann erfüllt, wenn die ϑ-adische Entwicklung der Zahl 1 abbricht oder periodisch wird[4]). Daraus ergibt sich

SATZ 7. *Ist die ϑ-adische Entwicklung der Zahl 1 periodisch, dann gilt für die Menge E aus Satz 6 die Relation*

$$\dim E = \frac{\log \vartheta}{\log g}.$$

8. Als weitere Anwendung unserer Sätze bestimmen wir die Hausdorffsche Dimension einiger Mengen, für die einzelne der Frequenzen der Ziffernkombinationen ij vorgeschrieben sind. Wir beschränken uns auf einen Spezialfall, aus dem aber klar ersichtlich ist, wie man bei speziellen Problemen vorzugehen hat.

SATZ 8. *Die Menge E aller reellen Zahlen $x \in [0, 1]$, in deren g-adischer Entwicklung die Ziffernpaare $(g-1, j_\nu)$, $(\nu = 1, 2, \ldots, m;\ 0 \leqq j_\nu \leqq g-1;\ j_m = g-1)$ mit den Frequenzen $p_{g-1, j_\nu} = a_\nu$ auftreten, hat die Hausdorffsche Dimension*

[4]) In diesem Fall gibt es für t_n nur endlich viele Werte, wenn mit t_n/ϑ^n die Länge eines ϑ-adischen Grundintervalles n-ter Stufe bezeichnet wird.

$\dim E = \frac{\log \vartheta}{\log g}$, *wo* ϑ *die eindeutig bestimmte Lösung* $\vartheta \geqq g-1$ *der Gleichung*

$$\text{(4a)} \qquad \vartheta^{a_m-2} - \left(\frac{1-\sum a_\nu}{g-m}\right)^{(1-\sum a_\nu)} \cdot \prod a_\nu^{a_\nu} \cdot \left(1-\frac{g-1}{\vartheta}\right)^{1-a_m} = 0$$

im Falle $1 \leqq m \leqq g-2$, *bzw.*

$$\text{(4b)} \qquad \vartheta^{a_m-2} - \prod a_\nu^{a_\nu} \cdot \left(1-\frac{g-1}{\vartheta}\right)^{1-a_m} = 0$$

für $m=g$ *bedeutet.* $(0^0=1)$.

BEWEIS. Wir bestimmen das Maximum der Funktion $\frac{-\sum\sum p_i p_{ij} \log p_{ij}}{\log g}$ unter den Nebenbedingungen $p_{g-1,j_\nu} = a_\nu$. Dann ergibt sich aus Satz 3 und Satz 5, daß die Hausdorffsche Dimension gleich diesem Maximum ist, wenn die Markowsche Kette mit den maximalen p_{ij} als Übergangswahrscheinlichkeiten ergodisch ist, was bei unserem Beispiel zutrifft. Zur Bestimmung des Maximums verwenden wir die Lagrangesche Multiplikatorenmethode. Diese ist zulässig, da aus dem Resultat ersichtlich ist, daß das Maximum im Innern des betrachteten Gebietes angenommen wird.

Die betrachtete Funktion ist $f = \sum_i \sum_j p_i p_{ij} \log p_{ij} = \sum_i p_i f_{p_i}$. Als Nebenbedingungen ergeben sich die Gleichungen

$$(\alpha) \qquad \sum_i p_i - 1 = 0$$

$$(\beta_i) \qquad \sum_j p_{ij} - 1 = 0 \qquad (i = 0, 1, \ldots, g-1)$$

$$(\gamma_j) \qquad \sum_i p_i p_{ij} - p_j = 0 \qquad (j = 0, 1, \ldots, g-1)$$

und

$$p_{g-1,j_\nu} = a_\nu, \qquad \nu = 1, 2, \ldots, m.$$

Wir müssen die Ableitungen der Funktion $f + \varkappa(\alpha) + \sum_i \mu_i(\beta_i) + \sum_j \lambda_j(\gamma_j)$ nach allen Variablen $p_i, p_{ij} \neq p_{g-1,j_\nu}$ bilden und gleich Null setzen. Dabei sind $\varkappa, \mu_i, \lambda_j$ die Lagrangeschen Multiplikatoren. Das ergibt die Gleichungen

$$(1)_i \qquad f_{p_i} + \varkappa + \sum_j \lambda_j(p_{ij} - \delta_{ij}) = 0 \qquad (i = 0, 1, \ldots, g-1)$$

$$(2)_{ij} \qquad p_i(1 + \log p_{ij}) + \mu_i + \lambda_j p_i = 0 \qquad (i,j) \neq (g-1, j_\nu).$$

Aus $(2)_{ij}$ und $(2)_{ik}$ ergibt sich $\log p_{ij} - \log p_{ik} = \lambda_k - \lambda_j$ für alle (i,j) und $(i,k) \neq (g-1, j_\nu)$. Für $i = 0, 1, \ldots, g-2$ gilt $1 = \sum_j p_{ij} = p_{i0} \sum_j e^{\lambda_0 - \lambda_j}$ wegen (β_i), und daraus folgt $p_{ij} = p_{lj}$ für $i, l \leqq g-2$ und alle j. Subtrahiert man $(1)_l$ von $(1)_i$, so erhält man

$$f_{p_i} - f_{p_l} + \varkappa - \varkappa + \sum_j \lambda_j(p_{ij} - p_{lj} + \delta_{lj} - \delta_{ij}) = 0$$

und daraus $\lambda_i = \lambda_l$ für alle $i, l \leqq g-2$, da $f_{p_i} = f_{p_l}$ ist. Also gilt $p_{ij} = p_{lk}$ für alle $i, j, l, k \leqq g-2$. Aus $(2)_{00}$ und $(2)_{0,g-1}$ ergibt sich $\lambda_{g-1} = \lambda_0 + \log \frac{p_{00}}{p_{0,g-1}}$.

Daraus folgt aus $(1)_0$ die Relation $\varkappa = -\log p_{00}$. Schließlich bekommt man aus $(1)_{g-1}$ die Gleichung

$$(4)\qquad \begin{cases} (1-\sum a_\nu)\log\dfrac{1-\sum a_\nu}{g-m} + \sum a_\nu \log a_\nu + \\ \quad + (a_m - 2)\log p_{00} - (a_m - 1)\log\big(1-(g-1)\,p_{00}\big) = 0, \end{cases}$$

wenn $m \leqq g-2$ ist. Für $m=g$ ($m=g-1$ ist unmöglich) fällt das Glied $(1-\sum a_\nu)\log\dfrac{1-\sum a_\nu}{g-m}$ weg.

Wir bekommen auf diese Art eine Gleichung für p_{00}. Nun folgt aber aus

$$\sum_i p_i \cdot (1)_i = \sum_i p_i f_{p_i} + \varkappa \sum_i p_i + \sum_i \sum_j \lambda_j (p_i p_{ij} - p_i \delta_{ij}) = 0$$

wegen (α), (γ_j) die Beziehung

$$f + \varkappa + \sum_j \lambda_j \Big(\sum_i p_i p_{ij} - p_j\Big) = f + \varkappa = 0$$

und daher $f = -\varkappa = \log p_{00}$. Daraus ergibt sich nun unmittelbar die Behauptung des Satzes, wenn noch gezeigt werden kann, daß es genau eine Zahl p_{00} mit $0 < p_{00} < \dfrac{1}{g-1}$ gibt, die (4) erfüllt. Nun ist diese Gleichung von der Form $\varphi(x) = x^\alpha - \gamma\big(1-(g-1)\,x\big)^\beta = 0$ mit $\alpha \geqq 1$, $\beta \geqq 0$, $\gamma > 0$. ($x = p_{00}$, für p_{00} muß gelten $(g-1)\,p_{00} \leqq 1$). Da $\varphi(0) = -\gamma < 0$ und $\varphi\left(\dfrac{1}{g-1}\right) = \left(\dfrac{1}{g-1}\right)^\alpha > 0$ ist und die Ableitung $\varphi'(x) = \alpha\, x^{\alpha-1} + \gamma\beta\big(1-(g-1)\,x\big)^{\beta-1} > 0$ ist, folgt das sofort.

9. Fordert man speziell $p_{g-1,g-m} = \cdots = p_{g-1,g-1} = 0$, $m \leqq g-2$, so reduziert sich Gl. (4a) auf die Gleichung

$$\vartheta^{-2} - \frac{1}{g-m}\left(1 - \frac{g-1}{\vartheta}\right) \quad \text{oder} \quad \vartheta^2 - (g-1)\,\vartheta - (g-m) = 0.$$

Man überzeugt sich leicht, daß die Matrix der p_{ij} mit der entsprechenden Matrix der Übergangswahrscheinlichkeiten bei der ϑ-adischen Entwicklung übereinstimmt.

Es gilt sogar ein allgemeiner Satz, der einen Zusammenhang mit ϑ-adischen Entwicklungen herstellt. Bevor wir ihn formulieren, wollen wir noch einen neuen Begriff einführen. Wir nennen eine Ziffernsequenz $a_1\, a_2 \ldots a_k$ „kleiner" als $b_1\, b_2 \ldots b_k$, wenn $a_1\, a_2 \ldots a_k$ vor $b_1\, b_2 \ldots b_k$ ist bei lexikographischer Anordnung.

Eine Menge von Sequenzen (der Länge k) heiße $[a_1\, a_2 \ldots a_k]$-Kette, wenn sie $a_1\, a_2 \ldots a_k$ und alle größeren Sequenzen enthält.

Es gilt nun der

SATZ 9. *Die Menge aller g-adischen Entwicklungen, die dadurch charakterisiert sind, daß in ihr die Sequenzen einer* $[g-1, a_2, \ldots, a_k]$*-Kette nicht vorkommen, besitzt die Hausdorffsche Dimension* $\dfrac{\log\vartheta}{\log g}$, *wo* $\vartheta > g-1$ *die eindeutig bestimmte Lösung der Gleichung*

$$(5)\qquad \vartheta^k - (g-1)\,\vartheta^{k-1} - \cdots - a_k = 0$$

ist.

BEWEIS. Sei $\vartheta^k=(g-1)\vartheta^{k-1}+\cdots+a_k$. Die Menge aller Folgen $\{x_\nu\}$, wo keine der Sequenzen einer $[g-1, a_2, \ldots, a_k]$-Kette auftritt, ist mit der Menge aller ϑ-zulässigen Folgen identisch[5]). Es gilt also nach Satz 7 (weil die Entwicklung $1=\frac{g-1}{\vartheta}+\frac{a_2}{\vartheta^2}+\cdots+\frac{a_k}{\vartheta^k}$ abbricht), daß $\dim E=\frac{\log\vartheta}{\log g}$ ist, wo ϑ Lösung von $x^k-(g-1)x^{k-1}-\cdots-a_k=0$ ist mit $\vartheta\geqq g-1$.

Wir müssen noch zeigen, daß ϑ durch die Bedingung $\vartheta>g-1$ wirklich eindeutig bestimmt ist. Sei $f(x)=x^k-(g-1)x^{k-1}-\cdots-a_k$. Dann ist $f(g-1)<0$. Sei nun $x>g-1$ eine beliebige Nullstelle von $f(x)$. Dann ist wegen

$$x^k=(g-1)x^{k-1}+\cdots+a_k$$

der Wert

$$\begin{aligned}xf'(x)&=k\left((g-1)x^{k-1}+\cdots+a_k\right)-\left((g-1)(k-1)x^{k-1}+\cdots+a_{k-1}x\right)\\&=(g-1)(k-k+1)x^{k-1}+a_2(k-k+2)x^{k-2}+\cdots+a_{k-1}(k-1)x+ka_k>0.\end{aligned}$$

Es ist also an jeder Nullstelle $f'(x)>0$. Gäbe es mehr als eine Nullstelle, so müßte an mindestens einer $f'(x)<0$ sein, was unserem Resultat widerspricht.

BEMERKUNG. Ähnliche Fragestellungen wie in Satz 9 wurden von VOLKMANN [6], [7] untersucht, wobei jedoch andere Methoden verwendet wurden. Der Spezialfall, daß die Ziffernsequenz $[g-1, g-1, \ldots, g-1]$ ausgeschlossen wird, ist in den Ergebnissen von VOLKMANN [6] enthalten.

Literatur

[1] BILLINGSLEY, P.: Hausdorff dimension in probability theory. Illinois J. Math. **4**, 187–209 (1960).
[2] CIGLER, J.: Ziffernverteilung in ϑ-adischen Brüchen. Math. Z. **75**, 8–13 (1961).
[3] EGGLESTON, H. G.: The fractional dimension of a set defined by decimal properties. Quart. J. Math. **20**, 31–36 (1949).
[4] KHINCHIN, A. I.: Mathematical foundations of information theory. New York: Dover 1957.
[5] KINNEY, J. R.: Singular functions associated with Markov chains. Proc. Amer. Math. Soc. **9**, 603–608 (1958).
[6] VOLKMANN, B.: Über Hausdorffsche Dimensionen von Mengen, die durch Ziffern-eigenschaften charakterisiert sind III. Math. Z. **59**, 259–270 (1953).
[7] VOLKMANN, B.: Über Hausdorffsche Dimensionen von Punktmengen, die durch Zifferneigenschaften charakterisiert sind V. Math. Z. **65**, 399–413 (1956).
[8] VOLKMANN, B.: Über Hausdorffsche Dimensionen von Mengen, die durch Ziffern-eigenschaften charakterisiert sind VI. Math. Z. **68**, 439–449 (1958).

Mainz, Mathematisches Institut der Universität

(Eingegangen am 30. Juli 1960)

[5]) Eine Folge ist genau dann ϑ-zulässig, wenn

$$\frac{x_{\nu+1}}{\vartheta}+\frac{x_{\nu+2}}{\vartheta^2}+\cdots+\frac{x_{\nu+k}}{\vartheta^k}<1 \quad \text{für alle } \nu, k \text{ gilt.}$$

RAYMOND, F.
Math. Zeitschr. 76, 31–41 (1961)

Local Cohomology Groups with Closed Supports

By
FRANK RAYMOND

1. Introduction

In the study of transformation groups certain theorems that hold for cohomology spheres have local analogues that hold for generalized manifolds (cohomology or homology manifolds). The Smith theorems are the classic examples. In [*1*] A. BOREL pointed out that if one takes the direct limit of the cohomology groups with *closed supports*, $H^p(U-x;K)$, where U is an open neighborhood of a given point x in a locally compact Hausdorff space X then the resulting limit group $I^p_x(X;K)$ possess the algebraic properties necessary for one to pass formally and more or less automatically from certain global situations to the analogous local situation. However one needs to know how this new local group corresponds to the standard local cohomology or local homology groups. In certain cases a correspondence is easy to establish, for example when X is a generalized manifold. In Chapter 13 of [*1*] BOREL used the correspondence in this particular case to localize his theorems on fixed points for elementary abelian groups. He also mentioned that using the correspondence which I shall establish in section 4 one should be able to deduce much of Chapter 5 from Chapter 4. In section 5 I shall show how one may deduce the standard local Smith theorems for generalized manifolds with comparatively simple formal manipulations from the corresponding Smith results for cohomology spheres.

The formula expressed as the exact sequence,

$$(4.3)\qquad \begin{cases} 0 \to \operatorname{Ext}\left(H^c_{p-1}(X, X-x;K), B\right) \to I^p_x(X;B) \\ \quad \to \operatorname{Hom}\left(H^c_{p+1}(X, X-x;K), B\right) \to 0, \end{cases}$$

establishes the desired correspondence for *clc* locally compact Hausdorff spaces where K denotes a principal ideal domain and B is a K-module. The correspondence between the somewhat analogous local homology "C"-groups and the standard local homology groups is much easier to derive (see [*8*] for definitions and correspondence). The difficulties in the case at hand arise partly from the lack of a relative sheaf cohomology theory.

In the first three sections the relative sheaf cohomology is developed so that in section 4 the different local groups may be compared. The remaining part of the paper is devoted to applications to transformation groups, generalized manifolds, and a new proof and generalization of SITNIKOV's duality theorem.

I would like to thank Professor BOREL for having called this problem to my attention and for several helpful conversations on the contents of this paper.

1. The cohomology sequence of a pair

Throughout the paper X will denote a topological space which satisfies at least the Hausdorff separation axiom and K will denote a principal ideal domain. If $\mathscr{S}$ is a sheaf on X it will usually mean a sheaf of K-modules. We shall follow the notation adopted in [5]. The restriction of a sheaf $\mathscr{S}$ on X to a subset A of X, the sections of S over A, and the canonical injective sheaf of cochains (see [2]) $\mathscr{C}^*(A; \mathscr{S}|A)$ are denoted by $\mathscr{S}|A$, $\mathscr{S}(A)$ and $\mathscr{C}(A; \mathscr{S})$, respectively. The sheaf on X which induces $\mathscr{S}|A$ on A and 0 on $X-A$ (A is locally closed) is denoted by $\mathscr{S}_A$. If Φ is a family of supports on X then $\Phi|A$ means those elements of Φ which are contained in A. When Φ is the family of all closed subsets of X then the sheaf cohomology of X with coefficients in the sheaf $\mathscr{S}$ will be denoted by $H^*(X; \mathscr{S})$ without any reference to the family Φ.

If A is a *closed* subset of X, $\mathscr{S}$ a sheaf on X, and Φ a family of supports (not necessarily paracompactifying) then there is an exact sequence of sheaves

$$0 \to \mathscr{S}_{X-A} \to \mathscr{S} \to \mathscr{S}_A \to 0 \tag{1.1}$$

which induces the exact cohomology sequence

$$\xrightarrow{\delta^*} H^p(X; \mathscr{S}_{X-A}) \to H^p(X; \mathscr{S}) \to H^p(X; \mathscr{S}_A) \to. \tag{1.2}$$

There is a natural homomorphism from $C^*_{\Phi|X-A}(X-A; \mathscr{S}|X-A)$ to $C^*_\Phi(X; \mathscr{S}_{X-A})$, which induces an isomorphism of the cohomology group $H^*_{\Phi|X-A}(X-A; \mathscr{S}|X-A)$ onto $H^*_\Phi(X; \mathscr{S}_{X-A})$ when Φ is paracompactifying, [5; II 4.10]. Then (1.2) becomes the exact sequence

$$\xrightarrow{\delta^*} H^*_{\Phi|X-A}(X-A; \mathscr{S}) \to H^p(X; \mathscr{S}) \to H^p_{\Phi|A}(A; \mathscr{S}) \to. \tag{1.3}$$

The cohomology module $H^*_\Phi(X; \mathscr{S}_{X-A})$ can be written in the classical way $H^*_\Phi(X, A; \mathscr{S})$. As we have seen, it is an invariant of $X-A$ and $\Phi|X-A$ when Φ is paracompactifying.

If U is an open subset of X the family of closed subsets of U is not in general a subfamily of a family of supports on X. The sheaf cohomology of U, $H^*(U; \mathscr{S})$, cannot be obtained in a natural way by concentrating the sheaf on U and taking cochains with supports in X as in the case of a closed subset. Also there doesn't seem to have been developed any corresponding relative theory as in the Čech and Alexander-Spanier cohomology theories. However, if one takes advantage of the fact that the sheaf of cochains $\mathscr{C}^*(X; \mathscr{S})$ is an injective sheaf, a suitable relative sheaf cohomology theory may be defined.

Since $\mathscr{C}^*(X; \mathscr{S})$ is flabby, $\Gamma\mathscr{C}^*(X; \mathscr{S}) \to \Gamma(\mathscr{C}^*(X; \mathscr{S})|U) \to 0$ is exact for any open subset U of X. This implies by [5; II Lemma 4.9.1] that

$$\mathscr{C}^*(X; \mathscr{S})|A = \mathscr{C}^*(A; \mathscr{S}|A) = \mathscr{C}^*(A; \mathscr{S}), \tag{1.4}$$

for any locally closed subset A of X, on which $\mathscr{S}$ is locally concentrated.

Let $\Gamma_{X,U}\mathscr{C}^*(X; \mathscr{S})$ be the kernel of the epimorphism $\Gamma\mathscr{C}^*(X; \mathscr{S}) \xrightarrow{i} \Gamma(\mathscr{C}^*(U; \mathscr{S})) \to 0$. Thus $\Gamma_{X,U}\mathscr{C}^*(X; \mathscr{S})$ *are the cochains of* X *which vanish*

on U; i.e., those cochains of X which have support on $X-U$. From the exact sequence of cochains

$$(1.5)\qquad 0\to\Gamma_{X,U}\left(\mathscr{C}^*(X;\mathscr{S})\right)\xrightarrow{j}\Gamma\left(\mathscr{C}^*(X;\mathscr{S})\right)\xrightarrow{i}\Gamma\left(\mathscr{C}^*(U;\mathscr{S})\right)\to 0,$$

one obtains the derived exact cohomology sequence,

$$(1.6)\qquad \to H^p(X,U;\mathscr{S})\xrightarrow{j^*}H^p(X;\mathscr{S})\xrightarrow{i^*}H^p(U;\mathscr{S})\xrightarrow{*}H^{p+1}(X,U;\mathscr{S})\to.$$

If V is an open subset of U then there exists a natural homomorphism of the exact sequence (1.6) for U into that for V. Now, if Y is an arbitrary subset of X the open sets U containing Y form an inverse system of sets directed by inclusion. Hence, we may form the direct limit of the exact sequences (1.6). If X is *metrizable* $\operatorname{Dir\,lim}_{U\supset Y}H^p(U;\mathscr{S})=H^p(Y;\mathscr{S})$ by [5; II 4.11.1]. Define $H^p(X,Y;\mathscr{S})=\operatorname{Dir\,lim}_{U\supset Y}H^p(X,U;\mathscr{S})$. Thus when X is metrizable the direct limit of the exact sequence (1.6) becomes the exact sequence

$$(1.7)\qquad \to H^p(X,Y;\mathscr{S})\xrightarrow{j^*}H^p(X;\mathscr{S})\xrightarrow{i^*}H^p(Y;\mathscr{S})\xrightarrow{d^*}H^{p+1}(X,Y;\mathscr{S})\to.$$

If $\mathscr{S}=K$, $H^p(X,Y;\mathscr{S})$ is isomorphic to the p-th relative Čech and p-th relative Alexander-Spanier cohomology group.

2. Characterization $H^*(H,U;\mathscr{S})$

In this section we define a module equivalent to $H^*(X,U;\mathscr{S})$ when X is locally compact and *hlc*. The method depends heavily on the work of Borel and Moore [2].

Let X be locally compact and $R(K)$ be an injective resolution of K induced by the imbedding of K into its field of quotients. Let the subscript c denote the family of *compact subsets* of a given topological space. Following Borel and Moore, define $\mathscr{C}_H(X;K)$ as the sheaf determined by the presheaf $U\to\operatorname{Hom}\left(\Gamma_c(\mathscr{C}^*(X;K)|U),R(K)\right)$. *The standard sheaf for homology of X with coefficients in the sheaf $\mathscr{S}$, $\mathscr{C}_H(X;\mathscr{S})$ is defined as $\mathscr{S}\otimes\mathscr{C}_H(X;K)$.* If Φ is any family of supports then $H_p^{\Phi}(X;\mathscr{S})=H_p\left(\Gamma_{\Phi}(\mathscr{C}_H(X;\mathscr{S}))\right)$.

In [2] it is shown that $\mathscr{C}_H(X;K)$ is a flabby sheaf, $\mathscr{C}_H(X;K)\,(U)$ is a torsion free graded differential module for every U open in X. Moreover, the natural homomorphism $\operatorname{Hom}\left(\Gamma_c(\mathscr{C}^*(X;K)|U),R(K)\right)\to\mathscr{C}_H(X;K)\,(U)$ is an isomorphism onto, and the homology theory is functorial with respect to proper maps. Hence, from the exact sequence of chains

$$(2.1)\quad 0\to\Gamma_c\left(\mathscr{C}_H(X;K)|U\right)\xrightarrow{i}\Gamma_c\left(\mathscr{C}_H(X;K)\right)\xrightarrow{j}\Gamma_c\mathscr{C}_H(X;K)/\Gamma_c\left(\mathscr{C}_H(X;K)|U\right)\to 0,$$

one obtains the exact derived homology sequence

$$(2.2)\qquad \xrightarrow{d_*}H_p^c(U;K)\xrightarrow{i_*}H_p^c(X;K)\xrightarrow{j_*}H_p^c(X,U;K)\xrightarrow{d_*}H_{p-1}^c(U;K)\to$$

The following lemma is taken from [2; 6.4, 6.5].

Lemma 1 (Borel-Moore). *Let B be a K-module and N an injective resolution of B finite degree. Let A be the sheaf defined by $U\to\operatorname{Hom}\left(\Gamma_c(\mathscr{C}_H(X;K)|U),N\right)$.*

Then

(i) $\mathscr{A}(U) = \mathrm{Hom}\left(\Gamma_c(\mathscr{C}_H(X; K)\, U), N\right)$,

(ii) *The sheaf* $\mathscr{A}$ *is bounded below;* $\mathscr{A}^p$ *is flabby for each* p,

(iii) *If* X *is hlc*^{r}, *then* $\mathscr{H}^p(\mathscr{A}) = 0$, $1 \leqq p \leqq r$, *and* $\mathscr{H}^0(\mathscr{A}) = B$,

(iv) *For each* U *open in* X, *the sequence*

$$(2.3)\quad 0 \to \mathrm{Ext}\left(H^c_{p-1}(U; K), B\right) \to H^p\left(\mathscr{A}(U)\right) \to \mathrm{Hom}\left(H^c_p(U; K), B\right) \to 0,$$

is exact and natural with respect to inclusion maps. In particular, if X *is hlc*^{r} *the module* $H^p\left(\mathscr{A}(U)\right) = H^p(U; B)$, $p \leqq r$ *(see the following discussion)*

Consider the commutative diagram

$$(2.4)\quad \begin{cases} 0 \to (\text{kernel } i_1) \to & \Gamma(\mathscr{A}) & \xrightarrow{i_1} & \Gamma(\mathscr{A}|U) & \to 0 \\ & \downarrow j & & \downarrow j & \\ 0 \to (\text{kernel } i_2) \to \Gamma\left(\mathscr{C}(X; \mathscr{A})\right) & & \xrightarrow{i_2} & \Gamma\left(\mathscr{C}(X; \mathscr{A})\, U\right) & \to 0. \end{cases}$$

The horizontal rows are exact because of (1.4) and the fact that the sheaf $\mathscr{A}$ and the sheaves of cochains are flabby. The top line of (2.4) is nothing but the dual of (2.1); thus kernel $i_1 = \mathrm{Hom}\left(\Gamma_c(\mathscr{C}_H(X; K))/\Gamma_c(\mathscr{C}_H(X; K)|U), N\right)$, i.e. the dual of the chains with compact support of X mod U. The terms of the second row may be regarded as double complexes since $\mathscr{A}$ is a graded differential sheaf. Then (2), (5) and (6) of [5; pp. 177–178] and the facts that $H^q(X; \mathscr{A}) = H^q(U; \mathscr{A}) = 0$ for $q > 0$ imply the natural homomorphisms j of the second and third columns induce isomorphisms onto in the derived cohomology diagram. The five lemma implies that H (kernel i_1) is mapped isomorphically onto the total cohomology of the double complex (kernel i_2). There is a spectral sequence $'E_r^{p,q}$ associated to the double complex $\Gamma\left(\mathscr{C}(X; \mathscr{A})\right)$, (see p. 177 of [5]), with the property $'E_1^{p,q} = C^p\left(X; H^q(\mathscr{A})\right)$ and $'E_2^{p,q} = H^p\left(X; H^q(\mathscr{A})\right)$. It can be seen that the spectral sequence associated to the subcomplex (kernel i_2) has $'E_1^{p,q}$ term equal to those p-cochains of X with coefficients in $H^q(\mathscr{A})$ whose support does not meet U. Thus the $'E_2^{p,q}$ term $= H^p\left(X, U; \mathscr{H}^q(\mathscr{A})\right)$. But by Lemma 1, $H^q(\mathscr{A}) = 0$ for $q \neq 0$, and B for $q = 0$. Thus the spectral sequence $'E_r^{p,q}$ is trivial and H^p (kernel i_2) $= H^p(X, U; B)$. Consequently we have

THEOREM 1. *If* X *is locally compact and hlc*^{r}, U *an open subset and* B *is a* K*-module, then the sequence*

$$(2.5)\quad 0 \to \mathrm{Ext}\left(H^c_{p-1}(X, U; K), B\right) \to H^p(X, U; B) \to \mathrm{Hom}\left(H^c_p(X, U; K), B\right) \to 0,$$

is exact and natural with respect to inclusion maps, for $0 \leqq p \leqq r$.

3. Excision

In this section it is shown that one may excise a closed subset from the relative sheaf cohomology.

We have seen that $H^p_\Phi(X, A; \mathscr{S})$ may be identified with $H^p_{\Phi|X-A}(X - A; \mathscr{S})$ when A is a closed subset of the Hausdorff space X, $\mathscr{S}$ a sheaf on X and Φ

is a paracompactifying family on X. Thus, if U is an open subset of X such that $U \subset A$, the $H^p_{\mathscr{P}}(X-U, A-U; \mathscr{S}) = H^p_{\mathscr{P}}(X, A; \mathscr{S})$. It will now be shown that an excision of a closed set F, $F \subset U$, for $H^p(X, U; \mathscr{S})$ is possible.

Consider the commutative diagram

$$(3.1)\quad \begin{cases} 0 \to & \Gamma_{X,U}(\mathscr{C}^*(X;\mathscr{S})) & \xrightarrow{j_1} & \Gamma(\mathscr{C}^*(X;\mathscr{S})) & \xrightarrow{i_1} & \Gamma(\mathscr{C}^*(U;\mathscr{S})) & \to 0 \\ & \downarrow i & & \downarrow i' & & \downarrow i'' & \\ 0 \to & \Gamma_{X-F,U-F}(\mathscr{C}^*(X-F;\mathscr{S})) & \xrightarrow{j_2} & \Gamma(\mathscr{C}^*(X-F;\mathscr{S})) & \xrightarrow{i_2} & \Gamma(\mathscr{C}^*(U-F;\mathscr{S})) & \to 0. \end{cases}$$

The epimorphisms i_1, i_2, i', and i'' are all induced by inclusions. The homomorphism i is well defined by $j_2^{-1} i' j_1$. It sends a cochain of X whose support is on $X-U$ by restriction into a class of cochains on $X-F$ with support on $X-U$. Clearly i has trivial kernel. It easily follows that i is also an epimorphism since i' is an epimorphism and $X-U$ is a closed subset of $X-F$. By passing to the derived of (3.1) one obtains

LEMMA 2 *(Excision for relative cohomology). If F and U are closed and open subsets of X, respectively, such that $F \subset U$, then the inclusion induces an isomorphism of $H^*(X, U; \mathscr{S})$ onto $H^*(X-F, U-F; \mathscr{S})$, for any sheaf $\mathscr{S}$ on X.*

From (1.7) we obtain the

COROLLARY. *If Y is an arbitrary subset of a metrizable space X and F is a closed subset of X such that $F \subset Y$, then the inclusion induces an isomorphism of $H^*(X, Y; \mathscr{S})$ onto $H^*(X-F, Y-F; \mathscr{S})$.*

4. Local cohomology groups

P. CONNER [4] has defined a type of local group which has been used by BOREL [1] for localizing global arguments in transformation group theory and by myself [9] for discussing compactifications of manifolds. The following is a modification of CONNER's cohomology of the ideal boundary.

If F is a subset of the space X, the family of open sets containing F form an inverse system of sets directed by inclusion. Let $i_{U,V}$ denote the inclusion map $V \subset U$.

4.1. DEFINITION. $I^p_F(X; \mathscr{S}) = \operatorname{Dir\,lim}\{H^p(U-F, \mathscr{S}), i^*_{U-F,V-F}\}$, where $F \subset V \subset U$.

LEMMA 3. *If X is paracompact and F is a closed subset or if X metrizable and F is an arbitrary subset then the sequence*

$$(4.2)\qquad \to H^p(X, X-F; \mathscr{S}) \xrightarrow{i^*} H^p(F; \mathscr{S}) \xrightarrow{i^*} I^p_F(X; \mathscr{S}) \xrightarrow{d^*},$$

is exact.

Proof. Consider the exact sequence

$$\xrightarrow{d^*} H^p(U, U-F; \mathscr{S}) \xrightarrow{i^*} H^p(U; \mathscr{S}) \xrightarrow{i^*} H^p(U-F; \mathscr{S}) \xrightarrow{d^*}.$$

Take the direct limit over all open sets U containing F and apply excision, (Lemma 2), to the first term, [5; II, Theorem 4.11.1] to the second term and the definition to the last term. This gives (4.2).

Note that when X is compact, F closed, and $\mathscr{S}$ is a K-module (4.2) is the same as [9; (2.6)], in fact, $I^p_F(X; B)$ is the same as what CONNER has called $I^p(X-F; B)$.

If F is a point (4.2) is exact without assuming X paracompact.

Let us denote the *reduced* cohomology groups by putting a tilda over the symbol for cohomology. *Then for* $x \in X$, *it is clear that* $\tilde{I}^p_x(X; B) = H^{p+1}(X, X-x; B)$ *if* $H^0(X, X-x; B) = 0$, and $I^p_x(X; B) = H^{p+1}(X, X-x; B)$ if $H^0(X, X-x; B) \neq 0$.

Combining (2.5) of Theorem 1 (where $U = X - x$) with the identification just mentioned above enables us to obtain our principal result on the local cohomology groups with closed supports viz,

THEOREM 2. *If* X *is locally compact and clc, then the sequence*

$$(4.3)\quad 0 \to \operatorname{Ext}\left(H^c_p(X, X-x; K), B\right) \to I^p_x(X; B) \to \operatorname{Hom}\left(H^c_{p+1}(X, X-x; B\right) \to 0$$

is exact, (with the obvious modifications in dimension 0).

We remark that if $K = B =$ *a* field then $\operatorname{Ext}\left(H^c_p(X, X-x; K), K\right) = 0$.

We are now interested in obtaining a converse to [1; Chapter 13, Lemma 3.2]. (This lemma also occurs independently in the proof of [9; Theorem 4].) As pointed out in [1], such a converse will be useful in transformation groups.

THEOREM 3. *Let* K *be a field,* X *be locally compact and clc over* K. *Then* X *is an* $n-gm$[1]) *over* K *(not necessarily orientable) if and only if* $I^p_x(X; K)$ $H^p(S^{n-1}; K)$ *for each* $x \in X$. *(S^{n-1} denotes the $(n-1)$-sphere).*

Proof. Since X is clc, the r-th dimensional local co-Betti number $p^r(x; X) = \dim H^c_r(X, X-x; K)$. By Theorem 2, $\tilde{I}^p_x(X; K) \approx H^c_{p+1}(X, X-x; K)$. Furthermore $H^c_0(X, X-x; K) = 0$. Thus the theorem follows.

It is also possible when K is the integers or a field to express in terms of the local group, $I^p_x(X; K)$, a condition implying that the $n-gm$ X be locally orientable. The proof uses the ideas exploited in [8] together with [7; Proposition] and the universal coefficient theorems. We shall state the result without proof as we shall not need it in what follows.

Let X *be an* $n-gm$ *over* K, *the integers or a field. Then* X *is locally orientable if and only if, for each* x, *there exists open sets* U, V *and* W $(x \in W \subset \overline{W} \subset V \subset U)$, *and a module* L *with* $K \cong L \subset H^{n-1}(V - \overline{W}; K)$, *such that* $i^*: H^r(U; K) \to H^r(V; K)$ *is trivial, for* $r \geq n-1$, *and the image*

$$\left(i^*_{U-y, V-\overline{W}}(H^{n-1}(U-y; K)) \subset H^{n-1}(V-\overline{W}; K)\right) = L,$$

for all $y \in W$.

5. Application to transformation groups

In this section we illustrate how Theorem 3 enables one to pass from a global situation to a local situation in transformation groups. We shall obtain the fixed point theorems for periodic mappings of generalized manifolds from the corresponding global theorem for cohomology spheres.

[1]) A generalized n-manifold or cohomology n-manifold (denoted by $n-gm$) is defined in [12, 1, or 10].

To use Theorem 3 one needs the fact that the direct limit of the spectral sequences of an inverse sequence of continuous maps is the spectral sequence of the limit map, (see [*1*; Chapter *5*]). This is true more generally for well ordered inverse systems but as we shall need open subsets to be paracompact we shall restrict ourselves to *locally compact metrizable spaces* for the duration of this section. We shall assume that the coefficient domain K and the transformation group G is isomorphic to the integers modulo a prime p, Z_p.

LEMMA 4 (SMITH). *Let Z_p act on the space X. If $\dim_{Z_p} X<\infty$ and is clc over Z_p, then the fixed point set F, $(F=F(Z_p, X))$ is clc over Z_p.*

A proof of this lemma may be found in [*1*; Chapter 4].

THEOREM 4 (SMITH). *Let X be a locally orientable $n-gm$ over Z_p. Suppose hat $Z_p=G$ acts on X. Then each component of $F(Z_p, X)$ is a locally orientable $-gm$.*

Proof. If a compact group G acts on a locally compact Hausdorf space X we may form the commutative diagram, (see [*1*] for further details).

$$\begin{cases} X \leftarrow X\times E \rightarrow E \\ \downarrow \qquad \downarrow \qquad \downarrow \\ X_G \xleftarrow{\pi_1} X\times_G E \xrightarrow{\pi_2} B_G. \end{cases} \tag{5.1}$$

The space E is acyclic up to a very large dimension N and is the total space for the N-universal classifying space B_G of the group G. The map π_2 is a fibering with fiber X and structural group G. If $G=Z_p$, as in the case at hand, the map π_1 over $X/G-F$ is a Vietoris map implying that $H^k(X_G; Z_p) \cong H^k((F\cap X)_G; Z_p)$ for $k>n$, $(X_G=X\times_G E,\ F_G=F\times_G E=F\times B_G)$. If Y is an invariant subset of X, the inclusion $i_{X,Y}: Y\subset X$ induces a homomorphism of the spectral sequence ${}_XE_r^{s,t}$ of the fibering $(X_G \xrightarrow{\pi_2} B_G)$ into the spectral sequence ${}_YE_r^{s,t}$ of the fibering $(Y_G \xrightarrow{\pi_2} B_G)$. Let $x\in F(Z_p, X)$. There exist a sequence $\{U_i\}$ of invariant neighborhoods which form a complete neighborhood system of x. For each U_i-x we have a diagram as above. By passing to a direct limit one obtains a spectral sequence $E_r^{s,t}=\operatorname{Dir\,lim}\{{}_iE_r^{s,t}\}$ of the inverse sequence of fiberings $\{(U_i-x)\times_G E \xrightarrow{\pi_2} B_G;\ i^*_{U_i-x,\,U_{i+1}-x}\}$. The E_2 term of this spectral sequence is

$$\begin{aligned} E_2^{s,t} &= \operatorname{Dir\,lim}\{H^s(B_G; H^t(U_i-x; Z_p))\} \\ &= H^s(B_G;\ I^t_x(X; Z_p) \approx H^s(B_G; H^t(S^{n-1}; Z_p)). \end{aligned}$$

Thus the local problem at x now coincides with the algebraic part of the argument used in [*1*; Chapter 4] to show that whenever Z_p acts on a cohomology sphere the fixed point set is a cohomology sphere. We may assume F is not o-dimensional at x for then x would be an isolated point. We have

$$\begin{aligned} \operatorname{Dir\,lim} H^k((U_i-x)_G; Z_p) &= \operatorname{Dir\,lim} \sum_{r+s=k} H^r((U_i-x)\cap F; Z_p)\otimes H^s(B_G; Z_p) \\ &= \operatorname{Dir\,lim} H^k(((U_i-x)\cap F)\times B_G; Z_p) \end{aligned}$$

for $k>n$. Since $E_2^{s,t}=E_\infty^{s,t}$ when F is not empty it follows $s-1<n-1$ from the equation above that $I_x^r(F;Z_p)\approx H^r(S^{s-1};Z_p)$ for some integer $s-1<n-1$. By Lemma 4, F is *clc* and therefore by (4.3) each point x has local homology groups

$$H_r^c(F,F-x;Z_p)=\begin{cases}0, & r\neq s\\ Z_p, & r=s.\end{cases}$$

However the integer s may depend upon the point x. We denote this dependence by $s(x)$. Clearly, if we show that $s(x)=s(y)$ for each x and y in a given component of F then Theorem 3 implies that this component is an $s-gm$ over Z_p. Local orientability will follow from the argument to be given.

Let us denote a component of F by the same letter, F. Choose U and V connected orientable invariant neighborhoods of $x\in F$ with the properties $\big(j^*(H_c^r(V;Z_p))\subset H_c^r(U;Z_p)\big)=0$, $r\neq n$, $\big(j^*(H_c^r(V\cap F;Z_p))\subset H_c^r(U\cap F;Z_p)\big)=0$, $r\neq s(x)$ and $\big(j^*(H_c^{s(x)}(W\cap F;Z_p))\subset H_c^{s(x)}(U\cap F;Z_p)\big)=\big(j^*(H_c^{s(x)}(V\cap F;Z_p))\subset H_c^{s(x)}(U\cap F;Z_p)\approx Z_p$, for any open invariant W, $x\in W\subset V$. Let $y\in V\cap F$. Choose V_y and U_y similar to the manner in which we choose V and U and so that $y\in V_y\subset U_y\subset V\subset U$. For each of the sets V_y, U_y, V and U consider the diagram (5.1) and the corresponding spectral sequences for the fiberings π_2 but now with *compact supports*. Because of the conditions on the choice of the sets V_y, U_y, V and U it is immediate that ${}_{V_y}E_\infty^{s,t}\to{}_{U_y}E_\infty^{s,t}$ and ${}_VE_\infty^{s,t}\to{}_UE_\infty^{s,t}$ are trivial when $t\neq n$. For $t=n$, ${}_{V_y}E_2^{s,n}={}_{V_y}E_\infty^{s,n}={}_{U_y}E_2^{s,n}={}_{U_y}E_\infty^{s,n}={}_VE_2^{s,n}={}_VE_\infty^{s,n}={}_UE_2^{s,n}={}_UE_\infty^{s,n}\approx Z_p$. The remarks about the diagram (5.1) imply the existence of the commutative diagram for all $k>n$,

$$\text{(5.2)}\qquad\begin{cases}0\to\sum\limits_{i+j=k}H_c^i(V_y\cap F;Z_p)\otimes H_c^j(B_G;Z_p)\to H_c^k\big((V_y)_G;Z_p\big)\\ \qquad\qquad\downarrow\qquad\qquad\qquad\qquad\downarrow\\ 0\to\sum\limits_{i+j=k}H_c^i(U_y\cap F;Z_p)\otimes H_c^j(B_G;Z_p)\to H_c^k\big((U_y)_G;Z_p\big)\\ \qquad\qquad\downarrow\qquad\qquad\qquad\qquad\downarrow\\ 0\to\sum\limits_{i+j=k}H_c^i(V\cap F;Z_p)\otimes H_c^j(B_G;Z_p)\to H_c^k(V_G;Z_p)\\ \qquad\qquad\downarrow\qquad\qquad\qquad\qquad\downarrow\\ 0\to\sum\limits_{i+j=k}H_c^i(U\cap F;Z_p)\otimes H_c^j(B_G;Z_p)\to H_c^k(U_G;Z_p).\end{cases}$$

Since $H^j(B_G;Z_p)\cong Z_p$, for all j it follows that $s(x)=s(y)$ for all $y\in V\cap F$, and that the neighborhood $V\cap F$ of X is an orientable neighborhood. This completes the proof.

Clearly, if F lies in an orientable part of X then F is orientable. Moreover, when $F\neq\varphi$, the local parity theorems follow immediately from the global parity theorems in the present framework.

6. Poincaré duality and Sitnikov duality

Let Φ be a paracompactifying family on a homology n-manifold X over K. Let $\mathscr{T}$ be the orientation sheaf $\mathscr{H}\big(\mathscr{C}_H(X;K)\big)$. Let $\mathscr{S}$ be a sheaf on X and let $\mathscr{A}^p=\mathscr{C}_H(X;\mathscr{S})_{n-p}$. Then $\mathscr{H}^0(\mathscr{A})=\mathscr{S}\otimes\mathscr{T}$, $H^p(\mathscr{A})=0$, $p>0$. Let F be

a closed subset of X then the diagram (Poincaré duality),

$$(6.1)\quad \begin{cases} \xrightarrow{i_*} & H_p^{\Phi}(X;\mathscr{S}) & \xrightarrow{j_*} H_p^{\Phi}(X,X-F,\mathscr{S}) \xrightarrow{d_*} & H_{p-1}^{\Phi|X-F}(X-F;\mathscr{S}) & \xrightarrow{i_*} \\ & \downarrow & & \downarrow & & \downarrow \\ \xrightarrow{j^*} H_{\Phi}^{n-p}(X;\mathscr{S}\otimes\mathscr{T}) & \xrightarrow{i^*} H_{\Phi|F}^{n-p}(F;\mathscr{S}\otimes\mathscr{T}) & \xrightarrow{d^*} H_{\Phi|X-F}^{n-p+1}(X-F;\mathscr{S}\otimes\mathscr{T}) \rightarrow \end{cases}$$

is commutative. The horizontal rows are exact and the vertical homomorphisms are all isomorphisms onto. The algebraic form of the Poincaré duality is proved in [3; Chapter 19]. This particular form for singular homology and for Čech homology with coefficients in a field is discussed in detail for example in the author's disertation. The recent work of BOREL-MOORE [2] now makes it possible to use the Borel-Moore homology theory.

We would now like to replace the role of the closed subset F in (6.1) by an arbitrary subset Y of X. Recall that if $S=K$ the sheaf $\mathscr{A}$ is a flabby resolution of the sheaf $\mathscr{T}$. Let U be an open subset of X. The same conditions that held for the sheaf (which was also called $\mathscr{A}$) in the proof of Theorem 1 now hold for the sheaf $\mathscr{A}$. Carrying through the formal algebraic analogue[1]) of (2.4) and (2.5) one obtains

THEOREM 5. *If X is a homology n-manifold then the diagram*

$$(6.2)\quad \begin{cases} \xrightarrow{i_*} & H_p(X;K) & \xrightarrow{j^*} & H_p(U;K) & \xrightarrow{d_*} H_{p-1}(X-U;K) & \xrightarrow{i_*} \\ & \downarrow & & \downarrow & \downarrow & \\ \xrightarrow{j^*} H^{n-p}(X;\mathscr{T}) & \xrightarrow{i^*} & H^{n-p}(U;\mathscr{T}) & \xrightarrow{d^*} & H^{n-p+1}(X,U;\mathscr{T}) & \xrightarrow{j^*} \end{cases}$$

is commutative. The horizontal rows are exact and the vertical homomorphisms are isomorphisms onto.

One should observe that the role of the closed and open subsets of X are interchanged in this new form of the Poincaré duality sequence.

Now let Y be an arbitrary subset of the homology n-manifold X and *assume* that X is *metrizable*. Taking the direct limit of (6.2) over all open U containing Y and using (1.7) one obtains the

COROLLARY. *Let Y be an arbitrary subset of the metrizable homology n-manifold X. Then the diagram*

$$(6.3)\quad \begin{cases} \xrightarrow{i_*} & H_p(X;K) & \xrightarrow{j_*} \operatorname{Dir\,lim}_{U\supset Y} H_p(U;K) \xrightarrow{d_*} & H_{p-1}^{\Phi|X-Y}(X-Y;K) & \xrightarrow{i_*} \\ & \downarrow & & \downarrow & \downarrow & \\ \xrightarrow{j^*} H^{n-p}(X;\mathscr{T}) & \xrightarrow{i^*} & H^{n-p}(Y;\mathscr{T}) & \xrightarrow{d^*} & H^{n-p+1}(X,Y;\mathscr{T}) & \xrightarrow{j^*} \end{cases}$$

is commutative, the rows are exact and the vertical homomorphisms are isomorphisms onto.

$$\left(H_{p-1}^{\Phi|X-Y}(X-Y;K) = \operatorname{Dir\,lim}_{U\supset Y} H_{p-1}(X-U;K)\right).$$

If H is an exact homology theory on the category of compact pairs then one may construct an exact homology theory on the category of arbitrary

[1]) The argument here is actually simpler since $\mathscr{A}$ is a resolution of $\mathscr{T}$.

Hausdorff pairs (X, A) by defining $H_p^c(X, A; K)$ to be the Dir $\lim_\lambda H_p(X_\lambda, A_\lambda; K)$ where (X_λ, A_λ) is the family of compact pairs in (X, A). (See [6; p. 255 Ex C—1].) In particular, if the homology n-manifold X in the corollary above is compact, the top row of (6.3) becomes

$$(6.4) \qquad \xrightarrow{i_*} H_p(X; K) \xrightarrow{j^*} H_p^c(X, X - Y, K) \xrightarrow{d_*} H_{p-1}^c(X - Y; K) \xrightarrow{i_*}.$$

This fact enables us to generalize and strengthen as well as to give a new simple proof of the *Sitnikov duality theorem* [11].

COROLLARY *(Sitnikov duality theorem). Let X be a metrizable homology n-manifold over K and Y an arbitrary subset of X, such that $H_p(X; K) = H_{p+1}(X; K) = 0$. If X is compact then $H_p^c(Y; K) = H^{n-p-1}(X - Y; \mathcal{F})$ and $H_p^c(X - Y; K) = H^{n-p-1}(Y; \mathcal{F})$. If X is not compact then in the duality the compact supports for the homology may be replaced by the families $\Phi|Y$ and $\Phi|X - Y$, respectively, where Φ is the family of closed subsets of X.*

If Y is an arbitrary finite dimensional separable metric space, then it may be embedded in some n-dimensional sphere for some n. SITNIKOV's proof of the corollary and the proof given above show that Sitnikov homology obtained from strengthened Vietories cycles is isomorphic to the direct limit of the Borel-Moore homology groups of the compact subsets of Y, $H_p^c(Y; K)$. We should also call attention to the Wilder-Kaplan form of the Poincaré duality for arbitrary subsets of an $n - gm$, see [12; ch. VIII, 9.1]. This differs from the Sitnikov form.

7. The Lefschetz duality theorem

For a relevant discussion of the properties of $n - gm$'s with boundaries the reader can refer to [10]. It is easy to see that if X is an $n - gm$ with boundary B over K then $H_p^c(X, X - x; K) = 0$, for all $x \in B$, all p. The Lefschetz duality theorem relates the cohomology of X with the homology of X mod B. It is usually assumed that there exists neighborhoods of B in X homeomorphic to $B \times$(unit interval). The next theorem shows that this assumption is unnecessary.

THEOREM 6. *Let X be an $n - gm$ with boundary B over K. Then $H_p^c(X; K) = H_p^c(X - B; K) \approx H_c^{n-p}(X - B; \mathcal{F})$ and $H^p(X; K) = H^p(X - B; K)$. Furthermore, if $X - B$ is orientable then $H^p(X; K) = H^p(X - B; K) \cong H_{n-p}(X, B; K) = H_{n-p}(X - B; K)$.*

Proof. The space X satisfies a Poincaré type of duality. In fact $H_p^c(X, X - B; K) = H_c^{n-p}(B; \mathcal{F})$. But $\mathcal{F}|B = 0$. Thus, $H_p^c(X, X - B; K) = 0$ for all integers p. Since X is *hlc*, $H^p(X, X - B; K) = 0$ for all integers p by Theorem 1. This implies that i^*: $H^p(X; K) \to H^p(X - B; K)$ and i_*: $H_p^c(X - B) \to H_p^c(X; K)$ are isomorphisms. The theorem now follows from the isomorphisms i_* and i^* and Poincaré duality.

We should point out that in the course of the proof we have shown that for an $n - gm$ X with boundary B, $I_B^p(X; K) = H^p(B; K)$. In addition, it follow that if X is paracompact $I_p(X - B; K) = H_{p-1}(B; K)$ where $I_p(X - B; K)$

is homology of the ideal boundary as defined in [9]. In particular, this answers a question I raised in [9; 6.2]. It also shows that the homology and cohomology of the boundary of an $n-gm\ X$ with boundary B is an invariant of $X-B$.

8. A problem of Conner

The groups $I_p(X;K)$ that were just mentioned in section 7 are the homology dual of the groups $I^p(X;K)$ defined by CONNER in [4]. It follows by [9; Theorem 2.16] that $I_p(X;K)=H^c_p(X', X'-\infty)$, $p>1$, where X' is the one point compactification of X by ∞. CONNER has asked what is the relationship between $I_p(X;K)$ and a group $I'_{p-1}(X;K)=\text{Inv lim}\{H^c_{p-1}(X-U_\lambda;K)\}$ where $\{U_\lambda\}$ is the family of open sets of X having compact closures. (In general the groups are not isomorphic.) From [9; Lemma 4.4] it follows that if X is clc and $\dim_K X<\infty$, then the one point compactification X' is clc if and only if $H^*_c(X;K)$ is finitely generated. In the Corollary of [8], it is shown that if a locally compact space X is clc at x, then there exists a natural isomorphism Ψ of $H^c_{p+1}(X, X-x;K)$ onto $\text{Inv lim}\{\tilde{H}^c_p(U_x-x;K)\}$ for each integer p, U_x is an open neighborhood of x. Thus one obtains the following answer to CONNER's question.

THEOREM 7. *If $\dim_K X<\infty$, X is clc and $H^*_c(X;K)$ is finitely generated, then $I_p(X;K)\cong I'_{p-1}(X;K)$.*

References

[1] BOREL, A.: Seminar on Transformation Groups. Annals of Mathematics Studies, No. 46, Princeton 1960.

[2] BOREL, A., and J. C. MOORE: Homology theory for locally compact spaces. Michigan Math. J. **7**, 137–160 (1960).

[3] CARTAN, H.: Séminaire Henri Cartan, 3e Année **1950–51**. Secrétariat mathématique, Paris 1955.

[4] CONNER, P.: On the impossibility of fibering certain manifolds by a compact fiber. Michigan Math. J. **4**, 249–256 (1957).

[5] GODEMENT, R.: Topologie algébrique et théorie des faisceaux. Paris: Hermann 1958.

[6] EILENBERG, S., and N. STEENROD: Foundations of Algebraic Topology. Princeton 1952.

[7] KWUN, K. W., and FRANK RAYMOND: Generalized cells in generalized manifolds. Proc. Amer. Math. Soc. **11**, 135–140 (1960).

[8] RAYMOND, FRANK: A note on the local "C" groups of GRIFFITHS. Michigan Math. J. **7**, 1–6 (1960).

[9] RAYMOND, FRANK: The end point compactification of manifolds. Pacific J. Math. **10**, 947–963 (1960).

[10] RAYMOND, FRANK: Separation and union theorems for generalized manifolds with boundary. Michigan Math. J. **7**, 7–21 (1960).

[11] SITNIKOV, K.: The duality law for non-closed sets. Doklady Akad. Nauk SSSR **81**, 359–362 (1951).

[12] WILDER, R. L.: Topology of Manifolds. Amer. Math. Soc. Coll. publication **32** (1949).

Institute for Advanced Study, Princeton, N.J. (U.S.A.)

(Eingegangen am 3. August 1960)

MARCUS, S.
Math. Zeitschr. 76, 42—45 (1961)

Les approximations diophantiennes et la catégorie de Baire

En hommage à Monsieur OSKAR PERRON, pour son 80ième anniversaire

Par

SOLOMON MARCUS

Il y a dans la théorie des approximations diophantiennes beaucoup de théorèmes concernant l'ensemble des nombres jouissant de telle ou telle propriété, qui utilisent la mesure ou le nombre cardinal. Je me propose, dans ce qui suit, de signaler deux théorèmes qui font intervenir la catégorie de Baire. Le théorème 1, quoique non explicité, résulte, sans difficulté, de la démonstration du théorème 22 de [6], si l'on tient compte que la partie commune d'une infinité dénombrable d'ensembles ouverts partout denses est un ensemble résiduel. Le théorème 2 peut être aussi établi par considérations de la théorie des nombres, comme nous l'indiquons ci-dessous. Mais nous voulons montrer — et c'est précisément le but de cette note — que ces théorèmes peuvent être déduits d'un résultat tout-à-fait étranger à la théorie des nombres, à savoir du

Théorème de M. K. FORT jr.: Si les points de discontinuité d'une fonction réelle, définie sur un intervalle réel quelconque, forment un ensemble partout dense sur cet intervalle, alors les points de dérivabilité de la fonction forment un ensemble de première catégorie. ([3]; voir aussi [2] et [8]).

THÉORÈME 1. *Soit f une fonction réelle, définie et positive pour $x>0$. Les nombres réels ξ tels que l'inégalité*

$$|\xi - p/q| < f(q) \tag{1}$$

ait seulement un système fini de solutions en nombres entiers p et q $(q>0)$, forment un ensemble de première catégorie.

DÉMONSTRATION. Soit φ une fonction positive pour $x>0$ et telle que

$$\lim_{x\to\infty} \frac{\varphi(x)}{f(x)} = 0. \tag{2}$$

Considérons la fonction F suivante: $F(x)=0$ si x est irrationnel, $F(p/q)=\varphi(q)$ si p et q sont des entiers positifs, premiers entre eux. F est discontinue en chaque point rationnel, donc, en vertu du théorème de M. K. FORT jr., ses points de dérivabilité forment un ensemble D de première catégorie. Désignons par N l'ensemble des nombres ξ tels que l'inégalité (1) ait seulement un

système fini de solutions en nombres entiers p et $q\,(q>0)$ et soit N_1 l'ensemble des points irrationnels de N. Nous allons montrer que $N_1 \subset D$. Soit, en effet, $\xi \in N_1$. Pour chaque suite $\{p_n/q_n\}$ de nombres rationnels tendant vers ξ il y a un entier positif n_1 tel qu'on ait

$$|\xi - p_n/q_n| \geqq f(q_n) \quad \text{pour} \quad n > n_1. \tag{3}$$

Pour x irrationnel, on a

$$\frac{F(x) - F(\xi)}{x - \xi} = 0. \tag{4}$$

D'autre part, en vertu de (3), on a, pour chaque $n > n_1$,

$$\left| \frac{F(p_n/q_n) - F(\xi)}{p_n/q_n - \xi} \right| \leqq \frac{\varphi(q_n)}{f(q_n)}$$

d'où, en tenant compte de (2), (4) et du fait que

$$\lim_{n \to \infty} q_n = +\infty$$

on déduit que F est dérivable en ξ $\left(F'(\xi)=0\right)$, donc $N_1 \subset D$. $N - N_1$ étant dénombrable et D étant de première catégorie, il s'ensuit que N est de première catégorie.

Remarque 1. Dans le cas particulier $f(x) = 1/x^{\varepsilon}$, avec $\varepsilon > 2$, l'ensemble N contient tous les nombres irrationnels algébriques, en vertu d'un théorème de K. F. Roth de [*10*].

Remarque 2. Il y a des cas où le complémentaire de N est de mesure nulle. Tel est le cas, par exemple, si $x^2 f(x)$ est continue, positive et non croissante sur $(0, +\infty)$ et l'intégrale

$$\int_0^{+\infty} x f(x)\, dx$$

est convergente. Voir, pour cela, un théorème de [*5*].

Remarque 3. Soit φ une fonction positive pour $x > 0$ et telle que $\lim_{x \to \infty} \varphi(x) = 0$. De la démonstration du théorème 1 on déduit, en prenant $f(x) \equiv 1$, que les nombres ξ tels que

$$\lim_{p/q \to \xi} \frac{\varphi(q)}{p/q - \xi} = 0$$

forment un ensemble de première catégorie.

Remarque 4. D'après un théorème de Hurwitz, on sait que, pour chaque nombre irrationnel x, l'inégalité

$$|x - p/q| < \frac{1}{\sqrt{5}\, q^2}$$

est satisfaite pour une infinité de paires de nombres entiers p et q $(q>0)$; $\sqrt{5}$ est ici la meilleure valeur possible [*4*].

Désignons, avec O. PERRON, par $M(x)$ la borne supérieure des nombres c tels que l'inégalité

$$|x - p/q| < \frac{1}{c\,q^2}$$

ait une infinité de solutions en nombres entiers p et q $(q>0)$. D'après un théorème de O. PERRON, $M(x)$ est fini pour chaque nombre irrationnel x dont les éléments du développement en fraction continue forment un ensemble borné, tandis que, pour tout autre x, $M(x)=+\infty$ [9]. D'après un théorème de BOREL, l'ensemble des nombres dont les éléments du développement en fraction continue forment un ensemble borné, est de mesure nulle [1]. En combinant ces deux théorèmes, on obtient:

Les nombres réels x pour lesquels $M(x)$ est fini forment un ensemble de mesure nulle.

Je vais compléter ce résultat par son analogue descriptif:

THÉORÈME 2. *L'ensemble E des nombres réels x pour lesquels $M(x)$ est fini, est de première catégorie.*

DÉMONSTRATION. Le théorème 2 peut être déduit à l'aide du théorème de O. PERRON de [9] et de la démonstration du théorème de BOREL de [1]. Dans cette dernière, on montre, en fait, que l'ensemble E des nombres réels, dont les éléments du développement en fraction continue forment un ensemble borné, est du type borélien F_σ. Du fait que E est de mesure nulle, il s'ensuit que E est un ensemble frontière. Mais un ensemble frontière, du type F_σ, est de première catégorie (Voir, par exemple, [7], p. 49).

La catégorie de l'ensemble E peut être établie aussi à l'aide du théorème 1 ci-dessus. Désignons, pour cela, par E_n l'ensemble des nombres irrationnels x tels que $M(x)<n$. L'ensemble $E - \bigcup_{n=1}^{\infty} E_n$, étant formé de nombres rationnels, est dénombrable, donc tout revient à montrer que E_n est de première catégorie pour chaque n entier positif. Soit, pour cela, $\xi \in E_n$. Il s'ensuit que l'inégalité

$$|\xi - p/q| < \frac{1}{n\,q^2}$$

a seulement un nombre fini de solutions en nombres entiers p et q $(q>0)$. Faisant usage du théorème 1, avec $f(x)=1/n\,x^2$, on déduit que E_n est de première catégorie.

REMARQUE 5. L'ensemble E a la puissance du continu. On a, par exemple, le résultat suivant de O. PERRON: Les nombres x tels que $M(x)=3$ forment un ensemble ayant la puissance du continu [9].

REMARQUE 6. Du théorème 1 et de la Remarque 2 on déduit que la catégorie de Baire, contrairement à la mesure de Lebesgue, est un instrument incapable de fournir une différentiation de nature globale, en ce qui concerne les divers modes d'approximation des nombres irrationnels par des rationnels. Ainsi, cette note, en attirant l'attention sur la catégorie de Baire, montre en même temps son inutilité dans la théorie des approximations diophantiennes, du moins pour le cercle de problèmes traités ci-dessus.

Mais la possibilité d'obtenir des résultats de la théorie des nombres à l'aide des théorèmes concernant la mesure, la catégorie, la puissance et le type borélien des ensembles liés à la dérivabilité des fonctions, mérite, sans doute, des recherches ultérieures.

Références

[1] Borel, E.: Les probabilités dénombrables et leurs applications arithmétiques. Rend. Circ. mat. Palermo **27**, 247–271 (1909).

[2] Ćetković, Simon: Un théorème de la théorie des fonctions. Comptes Rendus Acad. Sci. Paris **245**, 1692–1694 (1957).

[3] Fort jr., M. K.: A Theorem concerning functions discontinuous on a dense set. Amer. Math. Monthly **58**, 408–410 (1951).

[4] Hurwitz, A.: Über die angenäherte Darstellung der Irrationalzahlen durch rationale Brüche. Math. Annalen **39**, 279–284 (1891).

[5] Khintchine, A.: Einige Sätze über Kettenbrüche, mit Anwendungen auf die Theorie der Diophantischen Approximationen. Math. Annalen **92**, 115–125 (1924).

[6] Khintchine, A.: Fractions continues. [En russe.] Moscou 1949.

[7] Kuratowski, C.: Topologie. I. Warszawa-Wroclaw 1948.

[8] Marcus, S.: Points de discontinuité et points de dérivabilité. [En russe.], Revue Math. pures et appl. **2**, 471–473 (1957).

[9] Perron, Oskar: Über die Approximation irrationaler Zahlen durch rationale. I. S.-B. Heidelberg. Akad. Wiss. 1921, Abh. 4; II. S.-B. Heidelberg Akad. Wiss. 1921, Abh. 8.

[10] Roth, K. F.: Rational approximations to algebraic numbers. Mathematika. **2**, 1–20 (1955).

Institut de Mathématiques de l'Académie de la République Populaire Roumaine, Bucarest

(Eingegangen am 4. August 1960)

SCHMIDT, H.
Math. Zeitschr. 76, 46–50 (1961)

Über das Additionstheorem der zyklischen Funktionen

Von

HERMANN SCHMIDT

Die hier behandelte Frage nach der allgemeinen Lösung eines gewissen Systems von Funktionalgleichungen, dem die sogenannten „zyklischen Funktionen" genügen (vgl. unten (3), (6)) wurde von Herrn GHEORGHIU aufgeworfen, mit dem ein Briefwechsel stattgefunden hat. Wir benützen geläufige Methoden der Theorie der linearen Differentialsysteme in Matrixschreibweise, wobei ein Satz von POLYA [6] eine entscheidende Rolle spielt. In einem Zusatz gibt Herr GHEORGHIU eine Lösung für den Fall $n=3$ mittels eines von ROSÇULEȚ betrachteten hyperkomplexen Systems (der enge Zusammenhang mit den Matrizenmethoden liegt auf der Hand) unter Berufung auf einen Satz von SIERPINSKI [9].

Es sei $A=(a_{\nu\lambda})$ eine n-reihige quadratische Matrix mit Elementen aus einem Körper K. Eine solche Matrix heißt zyklisch, wenn für

$$\nu=1, 2, \ldots n-1, \lambda=1, 2, \ldots, n$$

(1) $$a_{\nu+1,\lambda}=a_{\nu,\lambda-1};$$

dabei sind alle Zeiger auf kleinste positive Reste mod n zurückgeführt zu denken. Eine zyklische Matrix ist demnach durch ihre erste Zeile $(a_{1\lambda})=(a_\lambda)$ vermöge

(2) $$a_{\nu\lambda}=a_{\lambda-\nu+1}$$

festgelegt, woraus nebenbei auch die Gültigkeit von (1) für $\nu=n$ folgt. Bezeichnet insbesondere D die zyklische Permutationsmatrix $D=(\delta_{\nu,\lambda-1})$ (Kroneckersches Symbol; Indizes stets mod n), so drückt sich (1) aus durch

(1′) $$DA=AD$$

und (2) durch die (zeilenweise zu verstehende) Beziehung

(2′) $$(a_{\nu\lambda})=(a_\varkappa)\,D^{\nu-1} \qquad (\nu=1, 2, \ldots, n),$$

oder auch durch

(2″) $$A=\sum_{\varkappa=1}^{n} a_\varkappa D^{\varkappa-1}.$$

Somit bilden die zyklischen Matrizen der festen Zeilenzahl n mit Elementen aus K eine Algebra, den Gruppenring zur endlichen zyklischen Gruppe ihrer Basiselemente $D^{\varkappa-1}$ $(\varkappa=1, 2, \ldots n)$; sie enthält wegen (1′) mit jeder nichtsingulären Matrix auch deren Kehrwert. Ist K bewertet, etwa der Körper

der reellen oder komplexen Zahlen mit der üblichen Betragbewertung, so enthält sie mit jeder (elementweise) konvergenten Folge auch deren Grenzwert.

Es bedeute nun $(u_\nu) = u$ eine einzeilige Matrix (reell- oder komplexwertiger) Funktionen einer (reellen oder komplexen) Veränderlichen x; ist dann $u = (f_\nu(x))$ jene Lösung des linearen Differentialsystems

$$\frac{du}{dx} = uD, \tag{3}$$

die der Anfangsbedingung $f_\nu(0) = \delta_{1\nu}$ entspricht, so heißen (vgl. Rey Pastor [5]) die $f_\nu(x)$ die (hyperbolischen) zyklischen Funktionen. Für $n = 2$ erhält man ch x, sh x, für $n = 3$ die drei Appellschen ([1], [10]) Funktionen ap x, cp x, sp x. Die durch die Zeile $(f_\nu(x))$ nach (2) festgelegte zyklische Matrix $U(x)$ ist wegen (2') die durch $U(0) = B\ (= (\delta_{\nu\lambda}))$ festgelegte Integralmatrix des Systems (3) $\left(\text{Normallösung von } \frac{d}{dx} U = UD\right)$.

Ist jetzt A *beliebig konstant*, und $U(x)$ die Normallösung von

$$\frac{d}{dx} U = UA, \tag{4}$$

so gilt bekanntlich die Matrixgleichung

$$U(x+y) = U(x) \cdot U(y) = U(y) \cdot U(x). \tag{5}$$

(Für festes y steht rechts wie links ein Fundamentalsystem, und diese stimmen für $x = 0$, also überall überein; überdies sind U und A vertauschbar.) Für $A = D$ ergibt sich daraus (1. Zeile) das *Additionstheorem der zyklischen Funktionen*

$$f_\nu(x+y) = \sum_{\varkappa=1}^{n} f_\varkappa(x)\, f_{\nu-\varkappa+1}(y). \tag{6}$$

(6) ist aber auch dann erfüllt, wenn A eine *beliebige konstante zyklische* Matrix, und $(f_\nu(x))$ dann die 1. Zeile der Normallösung von (4) bedeutet. Denn die letztere ist

$$U = \sum_{m=0}^{\infty} \frac{A^m x^m}{m!} = e^{xA} = \prod_{\varkappa=1}^{n} \exp\left(a_\varkappa x D^{\varkappa-1}\right),$$

also nach den vorausgeschickten Bemerkungen über zyklische Matrizen mit A zyklisch (wie man übrigens auch ohne Benützung der Reihenentwicklung aus (1') und (2') schließen kann), so daß aus (5) auch jetzt (6) folgt, worin die speziellen Koeffizientenwerte in A nicht eingehen. Es gilt also

Satz 1. *Ist A eine beliebige zyklische Matrix mit reellen oder komplexen, konstanten Elementen, so erfüllt die der Anfangsbedingung $u_\nu(0) = \delta_{1\nu}$ $(\nu = 1, 2, \ldots, n)$ genügende Lösung von $\frac{du}{dx} = uA$ das Additionstheorem* (6).

Bis zu einem gewissen Grade kann man nun auch die Gesamtheit der Lösungen von (6) überblicken. Man hat

SATZ 2. *Erfüllt das (reell- oder komplexwertige) Funktionssystem* $f_\nu(x)$ $(\nu = 1, 2, \ldots, n;\ -\infty < x < +\infty)$ *die Funktionalgleichungen* (6), *ist ferner* $f_\nu(x)$ *bei* $x = 0$ *stetig, und die Determinante der zyklischen Matrix* $U = (f_{\lambda-\nu+1}(x))$ *daselbst* $\neq 0$, *so ist* $(f_\nu(x))$ *die Anfangszeile der Normallösung eines Systems* (4) *mit konstanter, zyklischer Koeffizientenmatrix.*

Dies folgt aus einem Satz von PÓLYA ([6], [8]). Durch Rechtsmultiplikation der Zeile (6) mit $D^{\mu-1}$ $(\mu = 1, 2, \ldots, n)$ ergibt sich nämlich wegen (2') und (1') sofort (5) für die jetzt zu betrachtende Matrix U. Also ist nach POLYA U Normallösung eines Systems (4) mit festen Koeffizienten, und wegen $A = U^{-1}\frac{dU}{dx}$ ist A wieder zyklisch, w.z.b.w.

Vermöge $u = e^{\alpha x}z$ und $\beta x = t$ geht übrigens (4) über in

$$\frac{dZ}{dx} = Z(A - \alpha E) : \beta \qquad (\beta \neq 0);$$

in A kann also $a_{\nu\nu} = 0$ $(\nu = 1, 2, \ldots, n)$ und ferner ein von Null verschiedenes Glied a_λ gleich 1 angenommen werden. Für $n = 2$ kann also $A = D$, für $n = 3$

$$A = \begin{pmatrix} 0 & 1 & \gamma \\ \gamma & 0 & 1 \\ 1 & \gamma & 0 \end{pmatrix}$$

angenommen werden. Die (im Sinne von Satz 2) allgemeine Lösung wird dann für $n = 2$ $e^{\alpha x}\operatorname{ch}\beta x$, $e^{\alpha x}\operatorname{sh}\beta x$, für $n = 3$ aber erhält man nur bei der speziellen Wahl $\gamma = 0$ die mit den Appellschen Funktionen ([1], [10]) gebildete Lösung $e^{\alpha x}$ $(\operatorname{ap}\beta x,\ \operatorname{cp}\beta x,\ \operatorname{sp}\beta x)$. Für $\gamma = -1$ geht z.B. (mit $\alpha = 0$, $\beta = 1$) hervor

$$(7) \qquad \begin{cases} 3f_1(x) = 1 + 2\cos\sqrt{3}\,x, \quad 3f_2(x) = 1 + \sqrt{3}\sin\sqrt{3}\,x - \cos\sqrt{3}\,x, \\ 3f_3(x) = 1 - \sqrt{3}\sin\sqrt{3}\,x - \cos\sqrt{3}\,x. \end{cases}$$

Hier kann man sich durch unmittelbare Rechnung von der Gültigkeit von (6) überzeugen.

Betrachtet man anstatt der zyklischen Matrizen die Matrizen der Form $'A = (\sigma_{\nu\lambda}\, a_{\lambda-\nu+1})$, wobei gesetzt werden soll

$$\left.\begin{array}{rl} \sigma_{\nu\lambda} = +1 & \text{für } 1 \leqq \nu \leqq \lambda \leqq n \\ -1 & \text{für } 1 \leqq \lambda < \nu \leqq n \end{array}\right\} \text{(Indizes außerhalb dieser Grenzen hier ausgeschlossen!)},$$

oder also

$$(8) \qquad 'A = \sum_{\varkappa=1}^{n} a_\varkappa\, 'D^{\varkappa-1} \quad \text{mit} \quad 'D = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ & & \ldots\ldots & & \\ -1 & 0 & 0 & \ldots & 0 \end{pmatrix},$$

so bilden sie wieder eine Algebra. Den hyperbolischen zyklischen Funktionen $f_\nu(x)$ entsprechen die zirkulären Funktionen:

$$g_\nu(x) = \delta^{-\nu+1} f_\nu(\delta x), \quad \delta = e^{\frac{\pi i}{n}} \qquad (\nu = 1, 2, \ldots n; n = 2\text{: Kreisfunktionen}),$$

die die Anfangszeile der Normallösung von $\frac{d}{dx} U = U \cdot {}'D$ bilden. Man sieht, daß bei Ersetzung von D durch $'D$ (1') und (2') gültig bleiben, daß die Funktionalgleichungen gelten

$$(9) \qquad g_\nu(x+y) = \sum_{\varkappa=1}^{\nu} g_\varkappa(x)\, g_{\nu-\varkappa+1}(y) - \sum_{\varkappa=\nu+1}^{n} g_\varkappa(x)\, g_{n+\nu-\varkappa+1}(y),$$

und wieder Satz 1 und Satz 2 entsprechende Sätze ausgesprochen werden können, die aus jenen durch die Ersetzung $A \mid {}'A$ und von (6) durch (9) hervorgehen.

Zusatz von OCTAVIAN EM. GHEORGHIU, *Timişoara*

Führen wir die hyperkomplexe Funktion Z der reellen Variablen x

$$Z(x) = f(x) + \Theta g(x) + \Theta^2 h(x); \quad \Theta^3 - 1 = 0$$

ein, so ist die obige Matrixgleichung (5) mit

$$U = \begin{pmatrix} f(x) & g(x) & h(x) \\ h(x) & f(x) & g(x) \\ g(x) & h(x) & f(x) \end{pmatrix}$$

gleichwertig mit der hyperkomplexen Funktionalgleichung

$$Z(x+y) = Z(x) \cdot Z(y),$$

welche nichts anderes als CAUCHYS Gleichung ist, und deren allgemeine Lösung nach SIERPINSKI [9] durch $Z(x) = e^{\alpha x} e^{\Theta \beta x} e^{\Theta^2 \gamma x}$ mit reellen willkürlichen Konstanten gegeben ist[1]). Die drei reellen Komponenten sind

$$(10) \quad f(x) = e^{\alpha x} P(\beta x, \gamma x); \quad g(x) = e^{\alpha x} Q(\beta x, \gamma x); \quad h(x) = e^{\alpha x} R(\beta x, \gamma x),$$

wo P, Q, R die Appellschen Funktionen [1] sind, welche auch DEVISME [3] und BOSÇULEŢ [7] benutzt haben. Die Appellschen Funktionen P, Q, R von zwei Argumenten drücken sich mit Hilfe der zyklischen Appellschen Funktionen eines Arguments $\operatorname{ap} x$, $\operatorname{cp} x$ und $\operatorname{sp} x$ folgendermaßen aus:

$$(11) \qquad \big(P(\beta x, \gamma x), Q(\beta x, \gamma x), R(\beta x, \gamma x)\big) = \big(V(\beta x)\, \widetilde{V}(\gamma x)\big)_{1.\,\text{Zeile}},$$

wo $V(x)$ die zyklische Matrix mit der Anfangszeile $(\operatorname{ap} x, \operatorname{cp} x, \operatorname{sp} x)$, $\widetilde{V}(x)$ deren Transponierte bedeutet.

Demnach ist die allgemeine Lösung des Systems (5) für $n = 3$ durch (10) gegeben, worin die Werte (11) einzutragen sind.

[1]) Vgl. auch NAGUMO, Jap. J. Math. **13** (1936), § 5, S. 72. (H. Sch).

Literatur

[1] APPELL, P.: Sur certaines fonctions analogues aux fonctions circulaires. Comptes Rendus Ac. Sci. Paris **84**, 1378 (1877).

[2] BABINI: Sobre los ceros de las funciones cíclicas. Publ. Fac. Ci. exact. Fis. Nat., Buenos Aires B Nr. **12** (1932).

[3] DEVISME, J.: Ann. Fac Sci. Toulouse (III) **25**, 143–237 (1933).

[4] KNOLL, F.: Die zyklischen Funktionen und damit zusammenhängenden lineare Operatoren. Deutsche Math. **1**, 156–162 (1936).

[5] REY PASTOR: Theoria de las fonctiones cíclicas. Boletin Sem. mat. Argentino. Buenos Aires Nr. **4** (1929).

[6] PÓLYA, G.: Über die Funktionalgleichung der Exponentialfunktion im Matrizenkalkül. Sitzgsber. preuß. Akad. Wiss. Berlin **1928**, 96–99.

[7] ROȘULEȚ, M.: Studii si cercetări matematice **5**, Nr. 3–4, 361–401 (1954) (teză prezentată 1949, București).

[8] SCHWERDTFEGER, H.: Les fonctions de matrices. Act. sci. et industr. **649**, § 8, S. 45 (1938).

[9] SIERPINSKI: Fundamenta Math. **1**, 116–122, 125–129 (1920).

[10] WALLNER, K.: Die Funktion $\sum_{i=0}^{\infty} \frac{x^{3i}}{(3i)!}$ und ihre Abgeleiteten. Programm Realschule Rothenburg o. d. T., Nürnberg, Bieling-Dietz, 1913 (mit Verzeichnis älterer Literatur).

Würzburg, Mathematisches Institut der Universität

(Eingegangen am 11. August 1960)

SANDERS, B. L.
Math. Zeitschr. 76, 51–55 (1961)

Concerning a Theorem of Ky Fan's

By
B. L. SANDERS[1]

§ 1. Introduction

DEFINITION 1. Let f be a mapping of the set A into itself. A subset Q of A is an *invariant subset of* A with respect to the mapping f provided $f(Q) = Q$. The set Q is the *maximal invariant subset of* A with respect to f provided Q is an invariant subset of A with respect to f and every invariant subset of A with respect to f is a subset of Q.

DEFINITION 2. The set $\{B, C\}$ is a *decomposition of* A provided $A = B \cup C$ and $B \cap C$ is empty. (Either B or C or both may be empty.)

DEFINITION 3. Let $A_1, \ldots, A_n$ $(n \geqq 2)$ be a finite number of infinite sets arranged in cyclic order $(A_{n+1} = A_1)$. For each $i = 1, \ldots, n$ let f_i be a one-to-one mapping of A_i into A_{i+1}. The sets A_i with the mappings f_i have *property* K provided there exists a decomposition $\{B_i, C_i\}$ of each A_i such that

$$f_i(B_i) = C_{i+1}, \quad i = 1, \ldots, n-1 \quad \text{and} \quad f_n(B_n) = C_1.$$

NOTATION. Let $F_i = f_n \cdots f_i$, F_1^k for k an integer greater than one denote F_1 followed by itself $k-1$ times, and F_1^0 be the identity.

KY FAN [2] has shown that any finite collection of infinite sets $A_1, \ldots, A_n$ arranged in cyclic order and one-to-one mappings f_i that take A_i into A_{i+1} have property K when n is even. A. FRODA [3] has pointed out that such is not in general the case when n is odd. The purpose of this paper is to discuss necessary and sufficient conditions for such sets and mappings to have property K when n is odd. To this end we state for reference purposes the following theorem which is embodied (though not stated explicity as a theorem) in the paper by KY FAN referred to above.

THEOREM 1. *Let $A_1, \ldots, A_n$ be a finite collection $(n \geqq 2)$ of infinite sets arranged in cyclic order. For each $i = 1, \ldots, n$ let f_i be a one-to-one mapping of A_i into A_{i+1}. The sets A_i and the mappings f_i have property K if and only if there exists a decomposition $\{B_1, C_1\}$ of A_1 such that*

$$(1) \quad \begin{cases} A_1 - B_1 = [F_n(A_n) - F_{n-1}(A_{n-1})] \cup [F_{n-2}(A_{n-2}) - F_{n-3}(A_{n-3})] \\ \qquad \cup \cdots \cup [F_2(A_2) - F_1(B_1)] \qquad \text{for } n \text{ even} \end{cases}$$

[1]) This paper was prepared under contract No. AF 49(638)-598 with the United States Air Force Office of Scientific Research. The author would like to express his gratitude to R. L. PLUNKETT, PAUL J. MCCARTHY, and J. W. ELLIS for their advice and encouragement in connection with this research.

and

$$(2)\qquad \begin{cases} A_1 - B_1 = [F_n(A_n) - F_{n-1}(A_{n-1})] \cup [F_{n-2}(A_{n-2}) - F_{n-3}(A_{n-3})] \\ \qquad \cup \cdots \cup [F_3(A_3) - F_2(A_2)] \cup F_1(B_1) \qquad \textit{for } n \textit{ odd.} \end{cases}$$

§2. Main Results

THEOREM 2. *If $n(\geqq 3)$ is an odd integer a necessary condition that the collection $A_1, \ldots, A_n$ of infinite sets and the one-to-one mappings f_i of A_i into A_{i+1}, $i = 1, \ldots, n$ have property K is that every subset Q of A_1 which is invariant with respect to the mapping F_1 has a decomposition $\{Q', Q''\}$ such that*

$$(3)\qquad F_1(Q') = Q''.$$

PROOF. Let Q be any subset of A_1 which is invariant with respect to F_1. If Q is empty ($Q = \emptyset$), the decomposition $\{\emptyset, \emptyset\}$ satisfies (3). If Q is not empty, we define Q' and Q'' in the following manner. By hypothesis the sets $A_1, \ldots, A_n$ with the mappings $f_1, \ldots, f_n$ have property K. Thus by Theorem 1 there exists a decomposition $\{B_1, C_1\}$ of A_1 which satisfies (2). Let

$$(4)\qquad Q' = B_1 \cap Q \quad \text{and} \quad Q'' = C_1 \cap Q.$$

Clearly $\{Q', Q''\}$ is a decomposition of Q, and it is sufficient to show that (3) is satisfied.

Since $Q' \subset B_1$, $f_1(Q') \subset f_1(B_1) = C_2$. Thus $f_1(Q') \cap B_2 = \emptyset$. Hence $f_2 f_1(Q') \cap f_2(B_2) = \emptyset$, and since $f_2(B_2) = C_3$ and $\{B_3, C_3\}$ is a decomposition of A_3, we have that $f_2 f_1(Q') \subset B_3$. Since n is odd, we obtain, after a finite number of steps, that $F_1(Q') \subset C_1$. Since $Q' \subset Q$, $F_1(Q') \subset F_1(Q) = Q$. Hence

$$(5)\qquad F_1(Q') \subset [C_1 \cap Q] = Q''.$$

Similarly one may show that

$$(6)\qquad F_1(Q'') \subset Q'.$$

Since Q is not empty and $Q = Q' \cup Q''$, (5) and (6) are seen to imply that neither Q' nor Q'' is empty. We wish to show that

$$(7)\qquad F_1 \text{ maps } Q' \text{ onto } Q''.$$

To this end assume the contrary. Thus since $Q'' \neq \emptyset$, there exists an $x \in Q'' - F_1(Q')$. Since $F_1(Q) = Q$ and $x \in Q$, there exists a $y \in Q$ such that $F_1(y) = x$. Since $x \notin F_1(Q')$, $y \in Q''$. Thus $x = F_1(y) \in F_1(Q'')$. Hence by (6), $x \in Q'$. However, this is impossible since x was chosen in Q'' and $Q' \cap Q'' = \emptyset$. Thus (7) is verified and (5) and (7) establish (3).

REMARK. It is noted in [2] that

$$(8)\qquad A_1 \supset F_n(A_n) \supset F_{n-1}(A_{n-1}) \supset \cdots \supset F_1(A_1) \supset F_1 F_n(A_n) \supset \cdots$$

and is shown that in the case of even n, (1) is satisfied by $B_1 = S$ where

$$(9)\qquad \begin{cases} S = [A_1 - F_n(A_n)] \cup [F_{n-1}(A_{n-1}) - F_{n-2}(A_{n-2})] \\ \qquad \cup \cdots \cup [F_2(A_2) - F_1(A_1)] \cup \cdots. \end{cases}$$

The set S is the union of a denumerable collection of sets in which the k-th set is the complement of the $2k$-th set of (8) with respect to the $(2k-1)$-th set of (8).

THEOREM 3. *If $n(\geqq 3)$ is an odd integer a sufficient condition that the infinite sets $A_1, \ldots, A_n$ and the one-to-one mappings f_i of A_i into A_{i+1} have property K is that the maximal invariant subset Q of A_1 with respect to the mapping F_1 has a decomposition $\{Q', Q''\}$ such that*

$$F_1(Q') = Q''. \tag{10}$$

PROOF. It is easy to verify that the maximal invariant subset Q of A_1 with respect to the mapping F_1 is given by

$$Q = \bigcap_{k=0}^{\infty} F_1^k(A_1). \tag{11}$$

(Compare [1].) Either $Q = \emptyset$ or $Q \neq \emptyset$. Consider first the case in which $Q = \emptyset$. We define $B_1 = S$. (Note that this is exactly the same definition for B_1 as used in [2] for the case of even n.)

We wish to show that B_1 satisfies (2). We first show that the right-hand side of (2) is contained in the left-hand side of (2). Let x belong to the right-hand side of (2). Two possibilities exist: either

(i) $x \in [F_j(A_j) - F_{j-1}(A_{j-1})]$ where j is an odd integer such that $3 \leqq j \leqq n$,

or

(ii) $x \in F_1(B_1)$.

Consider case (i) first. Since j is odd, $j-1$ is even and from (9) we see that $H = [F_{j-1}(A_{j-1}) - F_{j-2}(A_{j-2})]$ is a term of B_1. By (i) $x \notin F_{j-1}(A_{j-1})$. Hence $x \notin H$. Further every term to the right of H in B_1 is by (8) a subset of $F_{j-1}(A_{j-1})$. Thus x does not belong to H nor to any term to the right of it in B_1. Consider the terms to the left of H in B_1. In each such term the set removed (that is, the second set) contains $F_j(A_j)$ and hence by (i) contains x. Thus x cannot belong to any term to the left of H in B_1. Thus x does not belong to B_1. By (i), $x \in A_1$, and thus $x \in A_1 - B_1$.

Consider next case (ii). From (9) we see that two possibilities exist: either

(I) $$x \in [F_1^k(A_1) - F_1^k F_n(A_n)], \text{ where } k \geqq 1$$

or

(II) $$x \in [F_1^k F_i(A_i) - F_1^k F_{i-1}(A_{i-1})], \text{ where } k \geqq 1 \text{ and } 2 \leqq i \leqq n.$$

In each case we look at the term of B_1 which gave rise to the term of $F_1(B_1)$ under consideration. Then we go to the $\left(\frac{n+1}{2}\right)$-th term of B_1 to the right of this term. This gives sets of the following types respectively:

(I') $$[F_1^k F_n(A_n) - F_1^k F_{n-1}(A_{n-1})],$$

(II') $$[F_1^k F_{i-1}(A_{i-1}) - F_1^k F_{i-2}(A_{i-2})] \quad \text{if } 2 < i \leqq n, \text{ and}$$
$$[F_1^k F_1(A_1) - F_1^k F_1 F_n(A_n)] \quad \text{if } i = 2.$$

In each case letting this final set correspond to H in case (i) and using an argument similar to the one there, we see that x cannot belong to this set nor to any term to the right or left of it in B_1. Hence in this case also $x \in A_1 - B_1$ and we have established that the right-hand side of (2) is contained in the left-hand side.

We now establish that the left-hand side of (2) is contained in the right-hand side. From (9) we obtain by appropriate set operations

$$(12) \quad \begin{cases} A_1 - B_1 = F_n(A_n) \cap \{[A_1 - F_{n-1}(A_{n-1})] \cup F_{n-2}(A_{n-2})\} \\ \qquad\qquad \cap \{[A_1 - F_{n-3}(A_{n-3})] \cup F_{n-4}(A_{n-4})\} \\ \qquad\qquad \cap \cdots \cap \{[A_1 - F_2(A_2)] \cup F_1(A_1)\} \cap \cdots. \end{cases}$$

Let $x \in A_1 - B_1$ and assume that x does not belong to the right-hand side of (2). Thus $x \notin [F_n(A_n) - F_{n-1}(A_{n-1})]$. By (12), $x \in F_n(A_n)$ and therefore $x \in F_{n-1}(A_{n-1})$. Hence $x \notin [A_1 - F_{n-1}(A_{n-1})]$, and from (12) we see that $x \in F_{n-2}(A_{n-2})$. By assumption $x \notin [F_{n-2}(A_{n-2}) - F_{n-3}(A_{n-3})]$, and hence $x \in F_{n-3}(A_{n-3})$. In a finite number of steps we obtain $x \in F_1(A_1)$. Suppose $x \in F_1^k(A_1)$ and k is even. Thus $x \notin [A_1 - F_1^k(A_1)]$ and since k is even, we see from (12) that $x \in F_1^k F_n(A_n)$. By (9), since $x \notin F_1(B_1)$, $x \notin [F_1^k F_n(A_n) - F_1^k F_{n-1}(A_{n-1})]$. Hence $x \in F_1^k F_{n-1}(A_{n-1})$. In a finite number of steps we obtain that $x \in F_1^{k+1}(A_1)$. If k is odd, we see from (9) that $x \notin [F_1^k(A_1) - F_1^k F_n(A_n)]$. Hence since $x \in F_1^k(A_1)$ by the inductive hypothesis, $x \in F_1^k F_n(A_n)$. As above a finite number of steps shows that $x \in F_1^{k+1}(A_1)$. Hence by induction $x \in F_1^k(A_1)$ for $k = 0, 1, 2, \ldots$. Thus from (11), $Q \neq \emptyset$. This is a contradiction. Hence x does belong to the right-hand side of (2) and the proof of this case is complete.

Consider now the case in which $Q \neq \emptyset$. Either $A_1 - Q = \emptyset$, or $A_1 - Q \neq \emptyset$. If $A_1 - Q = \emptyset$, define a decomposition $\{B_1, C_1\}$ of A_1 by $B_1 = Q'$, $C_1 = Q''$. Hence $A_1 - B_1 = A_1 - Q' = Q'' = F_1(Q') = F_1(B_1)$. Further, since $F_1(A_1) = A_1$, $F_j(A_j) = A_1$ for $1 \leqq j \leqq n$, and therefore $[F_j(A_j) - F_{j-1}(A_{j-1})] = \emptyset$ for all odd j such that $3 \leqq j \leqq n$, and (2) is verfied in this instance.

Now suppose that $A_1 - Q \neq \emptyset$. Let $x_1 \in A_1 - Q$. Define a set of points by letting $x_{i+1} = F_1^i(x_1)$ for $i = 1, 2, \ldots$. The points $x_1, x_2, \ldots$ thus obtained are distinct elements of $A_1 - Q$. Hence $A_1 - Q$ is infinite.

Define

$$(13) \quad \begin{cases} A_1' = A_1 - Q, \\ A_i' = A_i - f_{i-1} \cdots f_1(Q) \quad \text{for } i = 2, \ldots, n, \\ A_{n+1}' = A_1', \quad \text{and} \\ f_{i*} = f_i | A_i' \quad \text{for } i = 1, \ldots, n. \end{cases}$$

It is easy to verify that $A_1', \ldots, A_n'$ is an odd number of infinite sets arranged in cyclic order, that f_{i*} is a one-to-one mapping of A_i' into A_{i+1}' for $i = 1, \ldots, n$, and that the maximal invariant subset of A_1' with respect to $F_{1*} = f_{n*} \cdots f_{1*}$ is empty. Thus by the first case considered, the sets A_i' and the mappings f_{i*}

have property K. Hence by Theorem 1, there exists a decomposition $\{B_1', C_1'\}$ of A_1' such that

$$(14)\quad A_1' - B_1' = [F_{n*}(A_n') - F_{n-1*}(A_{n-1}')] \cup \cdots \cup [F_{3*}(A_3') - F_{2*}(A_2')] \cup F_{1*}(B_1').$$

Define a decomposition $\{B_1, C_1\}$ of A_1 by letting $B_1 = B_1' \cup Q'$ and $C_1 = C_1' \cup Q''$. Now $A_1 - B_1 = [A_1' - B_1'] \cup Q''$, and since $F_{j*}(A_j') = F_j(A_j')$ for $j = 1, \ldots, n$, we get from (14) that

$$(15)\quad A_1 - B_1 = [F_n(A_n') - F_{n-1}(A_{n-1}')] \cup \cdots \cup [F_3(A_3') - F_2(A_2')] \cup F_1(B_1') \cup Q''.$$

Now $F_1(B_1') \cup Q'' = F_1(B_1') \cup F_1(Q') = F_1(B_1' \cup Q') = F_1(B_1)$. Substituting this in (15), substituting for the A_i' in (15) from (13), and simplifying, we see that (2) is satisfied and the proof is complete.

§ 3. Conclusion

CALLAHAN and KNEALE [1] have shown for $n = 2$ that a set B_1 satisfies (1) if and only if $B_1 = S \cup Q$ where S is as given in (9) and Q is some subset of A_1 which is invariant with respect to the mapping $f_2 f_1$. It seems reasonable to conjecture that a similar result holds for any even n. In the light of the results of § 2, it also seems reasonable to conjecture that if n is odd, and the sets $A_1, \ldots, A_n$ with the mappings $f_1, \ldots, f_n$ have property K, then a set B_1 satisfies (2) if and only if $B_1 = S \cup Q'$ where S is as given in (9) and Q' is a subset of the maximal invariant subset Q of A_1 with respect to $f_n \cdots f_1$ such that $f_n \cdots f_1(Q') = Q - Q'$.

References

[1] CALLAHAN, F. P., and S. G. KNEALE: A note on the Schroeder-Bernstein theorem. Amer. Math. Monthly **64**, 423–424 (1957).

[2] FAN, KY: Note on a theorem of BANACH. Math. Z. **55**, 308–309 (1952).

[3] FRODA, A.: A review of [2]. Zbl. Math. **47**, 285 (1953).

Dept. of Math., Florida State Univ., Tallahassee, Fla. (U.S.A.)

(Eingegangen am 27. September 1960)

KUNZ, E.
Math. Zeitschr. 76, 56—74 (1961)

Differentialformen inseparabler algebraischer Funktionenkörper

Von
ERNST KUNZ

Die folgende Arbeit knüpft an (unveröffentlichte) Untersuchungen von F. K. SCHMIDT über Differentialformen algebraischer Funktionenkörper an. F. K. SCHMIDT wies darauf hin, daß es für die Behandlung inseparabler algebraischer Funktionenkörper K zweckmäßig sein wird, statt der üblicherweise benutzten Differentialformen von K über dem Konstantenkörper k solche von K über k^p zu betrachten, wenn $[k:k^p]$ endlich ist. Es soll also der Körper, über dem differenziert wird (der Differentialkonstantenkörper), verschieden vom arithmetischen Konstantenkörper k gewählt werden. Von dieser Idee wird im folgenden Gebrauch gemacht.

Es wird gezeigt werden, daß die Schwierigkeiten, auf die man stößt, wenn man versucht, die Theorie der Differentialformen algebraischer Funktionenkörper auf den Fall auszudehnen, daß der Konstantenkörper k unvollkommen ist, überwunden werden können, wenn man nur einen geeignet gewählten Differentialkonstantenkörper k_0 verwendet, wobei k_0 im allgemeinen von k verschieden ist. Für die ganze Arbeit wird vorausgesetzt, daß die Charakteristik p von k von Null verschieden ist, da sich neue Resultate im wesentlichen nur für unvollkommenes k ergeben werden. Die verwendeten Differentialkonstantenkörper k_0 sind dann immer Zwischenkörper von k^p und k, für die $[k:k_0]$ endlich ist. Ist k vollkommen, dann ist $k_0=k$. Hier erhält man aus den allgemeinen Sätzen die für vollkommenes k bekannten Sätze als Spezialfälle zurück. Ist $[k:k^p]$ endlich, so kann stets $k_0=k^p$ gewählt werden. Dies bedeutet, daß die verwendete Differentiation von K die absolute Differentiation ist (Differentiation über dem Primkörper), da bei jeder Differentiation von K der Körper k^p in Null abgebildet wird. Auch im allgemeinen Fall, wenn $[k:k^p]$ nicht endlich ist, handelt es sich bei der Differentiation über k_0 im Grunde um die absolute Differentiation, wobei nur noch zusätzlich geeignete Elemente in Null abgebildet werden, damit man zu einem endlichen Differentialmodul gelangt. Dabei ist es weitgehend willkürlich, welche Elemente von K in Null abgebildet werden, weshalb auch der Differentialkonstantenkörper nicht eindeutig bestimmt ist. Diese Nichteindeutigkeit des Differentialkonstanenkörpers spielt aber keine große Rolle, weil sich die wichtigsten, mit Hilfe der Differentialformen gewonnenen Begriffe, nämlich die kanonische Klasse eines abstrakten Modells V von K und das geometrische Geschlecht von V als unabhängig von der Wahl des Differentialkonstantenkörpers erweisen.

Bei der Darstellung des Verhaltens der Differentialformen auf abstrakten vollständigen Modellen V von K folge ich dem von ZARISKI in [7] für algebraisch abgeschlossenes k und projektive Varietäten V gegebenen Aufbau. Es zeigt sich, daß sich alle wesentlichen Sätze mühelos auf den Fall eines beliebigen Konstantenkörpers übertragen lassen. Im weiteren Verlauf der Arbeit wird die Änderung der kanonischen Klasse eines abstrakten vollständigen Modells von K bei algebraischer Erweiterung von K untersucht. Es stellt sich heraus, daß die entsprechenden Verhältnisse vorliegen, wie sie für algebraische Funktionenkörper einer Variablen bekannt sind.

Durch Anwendung auf algebraische Funktionenkörper K einer Variablen ergibt sich, daß die hier definierte kanonische Klasse des singularitätenfreien Modells V von K mit der im Riemann-Rochschen Satz auftretenden kanonischen Klasse übereinstimmt. Das geometrische Geschlecht von V erweist sich als identisch mit dem Geschlecht von K. Hierin hat man die hauptsächliche Rechtfertigung für die Einführung der Differentialkonstantenkörper zu erblicken. Bei Verwendung von Differentialformen über dem arithmetischen Konstantenkörper k gelangt man auf dem hier eingeschlagenen Weg nicht zur richtigen kanonischen Klasse.

In der Arbeit werden die Grundtatsachen über die Differentialmoduln und die Kählerschen Differenten als bekannt vorausgesetzt und oft stillschweigend benutzt. Sie sind z. B. in [4] zusammengestellt. Die Differenten werden je nach Bedarf als Ideale oder Divisoren aufgefaßt, ohne daß dies in der Schreibweise unterschieden wird.

Herrn Prof. F. K. SCHMIDT möchte ich für seine wertvollen Anregungen danken, ebenso den Herren Dr. R. BERGER und Dr. H. J. NASTOLD für ihre freundlichen Ratschläge bei der Abfassung dieser Arbeit.

§ 1. Charakterisierung regulärer geometrischer Stellenringe durch Differentiationen

Es sei k ein beliebiger Körper der Charakteristik $p > 0$ und $A = k[x_1, \ldots, x_n]$ eine nullteilerfreie affine k-Algebra. Für jedes Primideal $\mathfrak{p}$ von A sei $A_\mathfrak{p}$ der zugehörige Stellenring, dessen maximales Ideal ebenfalls mit $\mathfrak{p}$ bezeichnet werde. Das Ziel dieses Paragraphen ist der Beweis von

SATZ 1. *$A_\mathfrak{p}$ ist genau dann regulär, wenn sein Differentialmodul $M(A_\mathfrak{p})$ ein freier $A_\mathfrak{p}$-Modul ist.*

Ein Beweis für Ringe $A_\mathfrak{p}$ mit endlichem $M(A_\mathfrak{p})$ wurde in [4], S. 177, gegeben, ebenso das Analogon zu Satz 1 im Fall, daß Char. $k = 0$ ist. Der allgemeine Beweis verläuft im Grunde analog, nur daß einige technische Schwierigkeiten zu überwinden sind. Es müssen zunächst einige Vorbemerkungen vorausgeschickt werden, auf die auch später noch zurückgegriffen wird.

Es sei A als Polynomrestklassenring von der Form $A = k[X_1, \ldots, X_n]/\mathfrak{n}$ gegeben, wobei der Kern $\mathfrak{n}$ von den m Polynomen $F_1(X_1, \ldots, X_n), \ldots, F_m(X_1, \ldots, X_n)$ erzeugt werde. Ein Unterkörper k' von k, der sämtliche in den F_i $(i = 1, \ldots, m)$ auftretenden Koeffizienten enthält, heißt ein *Definitionskörper für A*. Bekanntlich gilt

HILFSSATZ 1. *$\mathfrak{n} \cap k'[X_1, \ldots, X_n]$ ist das von $F_1, \ldots, F_m$ in $k'[X_1, \ldots, X_n]$ erzeugte Ideal. Ist $\{x_\mu\}_{\mu \in M}$ eine Basis von k über k', so ist es auch eine Basis von A über $k'[x_1, \ldots, x_n]$: $A = \bigoplus_{\mu \in M} k'[x_1, \ldots, x_n] \cdot x_\mu$.*

Es werden später nur noch Definitionskörper k' mit den folgenden beiden Eigenschaften auftreten: 1. $k^p \subseteqq k'$, 2. $[k' : k^p]$ ist endlich. Wenn wir in Zukunft von Definitionskörpern sprechen, dann sind immer solche gemeint, welche die Eigenschaften 1. und 2. besitzen.

HILFSSATZ 2. *k' sei ein Definitionskörper von $A = k[x_1, \ldots, x_n]$, $\{\alpha_i\}_{i \in I}$ eine p-Basis von k über k' und es sei $k_0 = k^p[\{\alpha_i\}_{i \in I}]$. $\beta_1, \ldots, \beta_s$ sei eine p-Basis von k' über k^p. Dann gilt für den Differentialmodul $M(A)$ von A eine Zerlegung:*

$$M(A) = E \oplus F,$$

wobei E ein endlicher Untermodul von $M(A)$ ist, erzeugt von den Differentialen $dx_1, \ldots, dx_n$, $d\beta_1, \ldots, d\beta_s$, F ein freier Untermodul von $M(A)$ mit der Basis $\{d\alpha_i\}_{i \in I}$. E ist isomorph zum Differentialmodul $M(A/k_0)$ und F ist isomorph zu $A \otimes_{k_0} M(k_0/k^p)$, wo $M(k_0/k^p)$ den Differentialmodul von k_0 über k^p bedeutet.

BEWEIS. $\beta_1, \ldots, \beta_s, \{\alpha_i\}_{i \in I}$ ist eine p-Basis von k über k^p. Wir schreiben $A = k^p[\beta_1, \ldots, \beta_s, x_1, \ldots, x_n, \{\alpha_i\}_{i \in I}]$ als Polynomrestklassenring über k^p in der Form $A = k^p[Y_1, \ldots, Y_s, X_1, \ldots, X_n, \{Z_i\}_{i \in I}]/\mathfrak{m}$. Im Kern $\mathfrak{m}$ kommen sicher die Polynome $Y_j^p - \beta_j^p$ $(j = 1, \ldots, s)$ und $Z_i^p - \alpha_i^p$ $(i \in I)$ vor, ferner die Polynome $F_k(Y_1, \ldots, Y_s, X_1, \ldots, X_n)$ $(k = 1, \ldots, m)$, die man erhält, wenn man in den $F_k(X_1, \ldots, X_n)$ sämtliche Koeffizienten mit Hilfe der p-Basis $\beta_1, \ldots, \beta_s$ von k' über k^p ausdrückt und überall β_j durch Y_j ersetzt. Die genannten Polynome erzeugen aber auch $\mathfrak{m}$: Denn rechnet man zunächst modulo dem von den $Y_j^p - \beta_j^p$ $(j = 1, \ldots, s)$ und $Z_i^p - \alpha_i^p$ $(i \in I)$ erzeugten Ideal, so geht $k^p[Y_1, \ldots, Y_s, X_1, \ldots, X_n, \{Z_i\}]$ in $k[X_1, \ldots, X_n]$ über, die $F_k(Y_1, \ldots, Y_s, X_1, \ldots, X_n)$ dabei in $F_k(X_1, \ldots, X_n^q)$.

Für den Differentialmodul $M(A) = M(A/k^p)$ ergibt sich hieraus:

$$M(A) = \Big[\bigoplus_{j=1}^{s} A\, dY_j \oplus \bigoplus_{i=1}^{n} A\, dX_i \oplus \bigoplus_{i \in I} A\, dZ_i\Big]\Big/ M_0,$$

wo M_0 der von den Elementen $\sum_{j=1}^{s} \frac{\partial F_k}{\partial \beta_j}\, dY_j + \sum_{i=1}^{n} \frac{\partial F_k}{\partial x_i}\, dX_i$ $(k = 1, \ldots, m)$ erzeugte Untermodul von $\bigoplus_{j=1}^{s} A\, dY_j \oplus \bigoplus_{i=1}^{n} A\, dX_i \oplus \bigoplus_{i \in I} A\, dZ_i$ ist. Da die Erzeugenden von M_0 von dZ_i frei sind, findet man

$$M(A) = E \oplus F$$

mit

$$E = \Big[\bigoplus_{j=1}^{s} A\, dY_j \oplus \bigoplus_{i=1}^{n} A\, dX_i\Big]\Big/ M_0 = \sum_{j=1}^{s} A\, d\beta_j + \sum_{i=1}^{n} A\, dx_i$$

und

$$F = \bigoplus_{i \in I} A\, dZ_i = \bigoplus_{i \in I} A\, d\alpha_i.$$

Wegen $M(k_0/k^p) = \bigoplus_{i \in I} k_0\, d\alpha_i$ ist F isomorph zu $A \otimes_{k_0} M(k_0/k^p)$. Ferner ist $M(A/k_0) \cong M(A)/A\, dk_0 = M(A)/F \cong E$. Damit ist Hilfssatz 2 bewiesen.

FOLGERUNG. *Für jedes $A_\mathfrak{p}$ ist $M(A_\mathfrak{p}) = E_\mathfrak{p} \oplus F_\mathfrak{p}$, wo $E_\mathfrak{p}$ ein endlicher Modul isomorph zu $M(A_\mathfrak{p}/k_0)$ ist und $F_\mathfrak{p}$ der freie Modul $F_\mathfrak{p} = \bigoplus_{i \in I} A_\mathfrak{p}\, d\alpha_i \cong A_\mathfrak{p} \otimes_{k_0} M(k_0/k^p)$.*

Wir benötigen schließlich noch den körpertheoretischen

HILFSSATZ 3. *Es sei $K = k(x_1, \ldots, x_n)$ ein endlicher Erweiterungskörper von k mit $\dim \frac{K}{k} = r$. $\{y_\lambda\}_{\lambda \in \Lambda}$ sei eine p-Basis von k über k^p. Dann gibt es eine p-Basis von K über K^p von folgender Art: Sie besteht aus allen Elementen y_λ $(\lambda \in \Lambda)$ mit $t \leqq n$ Ausnahmen und $t + r$ weiteren Elementen aus K.*

BEWEIS. Es genügt, den Beweis für $n = 1$ zu führen.

1. Fall: Es sei $K = k(x)$ und x sei transzendent über k. Dann ist $\{y_\lambda\}_{\lambda \in \Lambda}$, x eine p-Basis von $k(x)$ über $k^p(x^p)$.

2. Fall: x sei separabel algebraisch über k. Dann ist $k(x) = k(x^p)$ und daher ist $\{y_\lambda\}_{\lambda \in \Lambda}$ auch eine p-Basis von $k(x)$ über $k^p(x^p)$.

3. Fall: x genüge einer irreduziblen Gleichung $x^{p^e} - \varkappa = 0$ $(\varkappa \in k)$. Wir betrachten zunächst den Fall $e = 1$: $x \notin k$, $x^p \in k$. Dann gilt auch $x^p \notin k^p$, $x^{p^2} \in k^p$. Es sei $x^p \in k^p[y_{\lambda_1}, \ldots, y_{\lambda_s}]$. Wegen $x^p \notin k^p$ kann ein geeignetes der p-Basiselemente, etwa y_{λ_1}, gegen x^p ausgetauscht werden, so daß $k^p[y_{\lambda_1}, \ldots, y_{\lambda_s}] = k^p[x^p, y_{\lambda_2}, \ldots, y_{\lambda_s}]$ wird. Es folgt, daß das System $\{y_\lambda\}_{\substack{\lambda \in \Lambda \\ \lambda \neq \lambda_1}}$, x^p ebenfalls eine p-Basis von k über k^p bildet. Da $1, x, \ldots, x^{p-1}$ eine Basis von $k(x)$ über k bilden, ergibt sich, daß das System $\{y_\lambda\}_{\substack{\lambda \in \Lambda \\ \lambda \neq \lambda_1}}$, x eine p-Basis von $k(x)$ über $k^p(x^p)$ ist. Für beliebiges e ergibt sich nun die Behauptung sofort durch vollständige Induktion, indem man sukzessive die Elemente $x^{p^{e-1}}$, $x^{p^{e-2}}$, $\ldots$, x zu k adjungiert.

Für eine beliebige algebraische Erweiterung $k(x)$ von k ergibt sich die Behauptung aus dem 2. und 3. Fall.

Wir sind jetzt in der Lage, Satz 1 zu beweisen. Wir wenden Hilfssatz 3 an auf den Quotientenkörper $K = k(x_1, \ldots, x_n)$ von $A_\mathfrak{p}$ und den Restklassenkörper $\mathfrak{K} = k(\dot{x}_1, \ldots, \dot{x}_n)$ von $A_\mathfrak{p}$. Als p-Basis von k über k^p verwenden wir das System $\{\alpha_i\}_{i \in I}$, $\beta_1, \ldots, \beta_s$ aus Hilfssatz 2. Nach Hilfssatz 3 dürfen wir annehmen, daß fast alle α_i sowohl in einer p-Basis von K über K^p als auch in einer p-Basis von $\mathfrak{K}$ über $\mathfrak{K}^p$ auftreten, etwa alle α_i mit $i \neq i_1, \ldots, i_t$. Setzt man $k_0' = k^p[\{\alpha_i\}_{i \neq i_1, \ldots, i_t}]$, so ist der p-Grad von k über k_0' gleich $s + t$. In der p-Basis von K über K^p treten daher außer den α_i mit $i \neq i_1, \ldots, i_t$ noch $s + t + r$ weitere Elemente auf, in der p-Basis von $\mathfrak{K}$ über $\mathfrak{K}^p$ noch $s + t + r'$ weitere Elemente $\left(r = \dim \frac{K}{k},\ r' = \dim \frac{\mathfrak{K}}{k}\right)$.

Nach der Folgerung aus Hilfssatz 2 ist $M(A_\mathfrak{p}) = \left[E_\mathfrak{p} \oplus \bigoplus_{i = i_1, \ldots, i_t} A_\mathfrak{p}\, d\alpha_i\right] \oplus \bigoplus_{\substack{i \in I \\ i \neq i_1, \ldots, i_t}} A_\mathfrak{p}\, d\alpha_i$, woraus unmittelbar $E_\mathfrak{p} \oplus \bigoplus_{i = i_1, \ldots, i_t} A_\mathfrak{p}\, d\alpha_i \cong M(A_\mathfrak{p}/k_0')$ folgt. $M(A_\mathfrak{p})$ ist daher genau dann ein freier $A_\mathfrak{p}$-Modul, wenn $M(A_\mathfrak{p}/k_0')$ es ist. $M(A_\mathfrak{p}/k_0')$

ist genau dann frei, wenn sein Rang (= minimale Erzeugendenanzahl) mit dem Rang von $M(K/k_0')$ übereinstimmt. Nach dem Rangsatz aus [4], S. 171, gilt

$$\text{Rang } M\left(\frac{A_\mathfrak{p}}{k_0'}\right) = \text{Rang } M\left(\frac{\mathfrak{R}}{k_0'}\right) + \text{Rang } (\mathfrak{p}/\mathfrak{q} + \mathfrak{p}^2) \quad \text{mit} \quad \mathfrak{q} = (\mathfrak{p} \cap k_0'[A_\mathfrak{p}^p])\, A_\mathfrak{p}.$$

$M(A_\mathfrak{p})$ ist somit genau dann frei, wenn

$$\text{(1)} \qquad \text{Rang } (\mathfrak{p}/\mathfrak{q} + \mathfrak{p}^2) = \text{Rang } M\left(\frac{K}{k_0'}\right) - \text{Rang } M\left(\frac{\mathfrak{R}}{k_0'}\right)$$

gilt.

Es sei nun π ein Element von $\mathfrak{p} \cap k_0'[A_\mathfrak{p}^p]$. Dann ist wegen $k_0'[A_\mathfrak{p}^p] = A_\mathfrak{p}^p[\{\alpha_i\}_{i \neq i_1, \ldots, i_t}]$ sicher $\pi = f(\{\alpha_i\}_{i \neq i_1, \ldots, i_t})$, wo f ein Polynom mit Koeffizienten aus $A_\mathfrak{p}^p$ ist, das in allen Variablen höchstens vom Grad $p-1$ ist. Modulo $\mathfrak{p}$ erhält man $0 = \dot{f}(\{\alpha_i\}_{i \neq i_1, \ldots, i_t})$, wo $\dot{f}$ aus f entsteht, indem man sämtliche Koeffizienten durch ihre Restklassen mod $\mathfrak{p}$ ersetzt. $A_\mathfrak{p}^p$ geht modulo $\mathfrak{p}$ in $\mathfrak{R}^p$ über, $\dot{f}$ hat daher Koeffizienten aus $\mathfrak{R}^p$. Nun sind die $\alpha_i (i \neq i_1, \ldots, i_t)$ aber nach Voraussetzung über $\mathfrak{R}^p$ p-unabhängig, mithin ist $\dot{f} \equiv 0$. Alle Koeffizienten von f liegen daher in $\mathfrak{p} \cap A_\mathfrak{p}^p \subseteq \mathfrak{p}^p$. Folglich ist $\pi \in \mathfrak{p}^p$ und daher $\mathfrak{q} \subseteq \mathfrak{p}^p$. $\mathfrak{q}$ kann also in Formel (1) weggelassen werden.

Wegen $k_0'(K^p) = K^p[k_0'] = K^p[\{\alpha_i\}_{i \neq i_1, \ldots, i_t}]$ ergibt sich nach dem weiter oben Gesagten $\text{Rang } M\left(\frac{K}{k_0'}\right) = p\text{-Grad } \frac{K}{k_0'(K^p)} = s + t + r$. Die entsprechende Gleichung gilt auch mit $\mathfrak{R}$ anstelle von K und r' anstelle von r. Trägt man in (1) ein, so erhält man die gleichwertige Beziehung

$$\text{(2)} \qquad \text{Rang } \mathfrak{p}/\mathfrak{p}^2 = r - r' = \dim \frac{K}{k} - \dim \frac{\mathfrak{R}}{k}.$$

Diese ist aber mit der Regularität von $A_\mathfrak{p}$ gleichbedeutend. Damit ist Satz 1 bewiesen.

§ 2. Differentialkonstantenkörper

Der Modul $M(A_\mathfrak{p})$ ist genau dann endlich erzeugt, wenn $[k : k^p]$ endlich ist. Bei unendlich erzeugtem $M(A_\mathfrak{p})$ wird es für das Folgende notwendig sein, dadurch zu einem endlichen Modul überzugehen, daß man einen uninteressanten direkten Summanden von $M(A_\mathfrak{p})$ abspaltet. Wir definieren daher:

DEFINITION. Ein Körper k_0 mit $k^p \subseteq k_0 \subseteq k$ heißt *Differentialkonstantenkörper für* $A_\mathfrak{p}$, wenn

1. $\langle A_\mathfrak{p}\, dk_0 \rangle$ ein direkter Summand von $M(A_\mathfrak{p})$ ist und $\langle A_\mathfrak{p}\, dk_0 \rangle \cong A_\mathfrak{p} \otimes_k M(k_0/k^p)$,

2. $M(A_\mathfrak{p}/k_0) = M(A_\mathfrak{p})/\langle A_\mathfrak{p}\, dk_0 \rangle$ endlich ist.

k_0 heißt *Differentialkonstantenkörper für die affine k-Algebra A*, wenn k_0 Differentialkonstantenkörper für jeden lokalen Ring $A_\mathfrak{p}$ von A ist.

Hilfssatz 2 und seine Folgerung zeigen, daß es für jede affine k-Algebra A einen Differentialkonstantenkörper gibt. Da nach 1. $\langle A_\mathfrak{p}\, dk_0 \rangle$ frei ist und direkter Summand von $M(A_\mathfrak{p})$ ergibt sich aus Satz 1 nun sofort

SATZ 2. *Ist k_0 ein Differentialkonstantenkörper für $A_\mathfrak{p}$, so ist $A_\mathfrak{p}$ dann und nur dann regulär, wenn $M(A_\mathfrak{p}/k_0)$ frei ist.*

Wir zeigen noch zwei Hilfssätze über Differentialkonstantenkörper:

HILFSSATZ 4. *Ist k_0 ein Differentialkonstantenkörper eines Stellenringes $A_\mathfrak{p}$, so ist $[k:k_0]$ endlich.*

BEWEIS. Für den Quotientenkörper $K=k(x_1, \ldots, x_n)$ von $A_\mathfrak{p}$ ergibt sich nach 2., daß $M(K/k_0)$ endlich ist, daher ist K über $k_0(K^p)$ endlich erzeugt. Wegen $K^p=k^p(x_1^p, \ldots, x_n^p)$ ist $k_0(K^p)=k_0(x_1^p, \ldots, x_n^p)$ und daher ist K über k_0 endlich erzeugt, folglich erst recht k über k_0.

HILFSSATZ 5. *Gegeben seien endlich viele nullteilerfreie affine k-Algebren $A_1, \ldots, A_s$ und ein Körper $\hat{k}$ mit $k^p \subseteqq \hat{k} \subseteqq k$ und $[k:\hat{k}]<\infty$. Dann gibt es stets einen gemeinsamen Differentialkonstantenkörper k_0 für $A_1, \ldots, A_s$ mit $k_0 \subseteqq \hat{k}$.*

BEWEIS. Es sei $\{x_\lambda\}_{\lambda \in \Lambda}$ eine p-Basis von $\hat{k}$ über k^p und es sei $k=\hat{k}[\beta_1, \ldots, \beta_t]$. k' sei ein gemeinsamer Definitionskörper für $A_1, \ldots, A_s$, der außerdem noch $\beta_1, \ldots, \beta_t$ enthalte. Dann ist $k'[\{x_\lambda\}_{\lambda \in \Lambda}]=k$. Aus dem System $\{x_\lambda\}_{\lambda \in \Lambda}$ kann eine p-Basis $\{x_{\lambda'}\}_{\lambda' \in \Lambda'}$ von k über k' ausgewählt werden. Der Körper $k_0=k^p[\{x_{\lambda'}\}_{\lambda' \in \Lambda'}]$ hat dann nach Hilfssatz 2 alle gewünschten Eigenschaften.

§ 3. Differentialformen auf abstrakten Varietäten

Es sei $K=k(x_1, \ldots, x_n)$ ein algebraischer Funktionenkörper der Dimension r über k und V ein *vollständiges abstraktes Modell* von K/k, d.h. 1. V besteht aus allen lokalen Ringen von Primidealen von endlich vielen affinen k-Algebren $A_1, \ldots, A_s$ mit $Q(A_i)=K (i=1, \ldots, s)$; 2. Ist R ein beliebiger Bewertungsring von K/k mit dem maximalen Ideal $\mathfrak{p}$, so gibt es genau einen lokalen Ring $\mathfrak{o}$ aus V, der von R dominiert wird, d.h. für genau ein $\mathfrak{o} \in V$ gilt $\mathfrak{o} \subseteqq R$ und $\mathfrak{m}=\mathfrak{p} \cap \mathfrak{o}$, wenn $\mathfrak{m}$ das maximale Ideal von $\mathfrak{o}$ bedeutet. Der Ring $\mathfrak{o}$ heißt dann das Zentrum von R auf V.

DEFINITION. Es sei R ein Bewertungsring von K/k und $\mathfrak{p}$ sein maximales Ideal. $\mathfrak{p}$ heißt *Primdivisor 1. Art auf V*, wenn das Zentrum von R auf V die Restklassendimension $r-1$ hat*). Unter der *Divisorengruppe von V* verstehen wir die freie multiplikative Gruppe, die von den Primdivisoren 1. Art auf V erzeugt wird.

Das Zentrum $\mathfrak{o}$ von R auf V ist der lokale Ring eines minimalen Primideals eines Ringes A_i. Es sei A_i' die ganze Abschließung von A_i in K. Dann folgt: R ist der lokale Ring eines minimalen Primideals von A_i', also ein diskreter Bewertungsring.

DEFINITION. Ein Unterkörper k_0 von k heißt ein *Differentialkonstantenkörper für V*, wenn k_0 Differentialkonstantenkörper für alle lokalen Ringe von V und für alle Bewertungsringe von Primdivisoren 1. Art auf V ist.

Aus Hilfssatz 5 folgt sofort, daß V immer einen Differentialkonstantenkörper besitzt, weil alle in der Definition genannten Ringe aus endlich vielen

*) d.h. der Restklassenkörper des Zentrums soll über k den Transzendenzgrad $r-1$ haben.

affinen k-Algebren hervorgehen. Die Differentialmoduln $M(R/k_0)$ und $M(\mathfrak{o}/k_0)$ der Bewertungsringe R von Primdivisoren 1. Art auf V und der regulären lokalen Ringe $\mathfrak{o} \in V$ sind dann endliche freie Moduln. Der Differentialmodul $M(K/k_0)$ des Körpers K entsteht aus diesen Moduln, indem man sie tensoriell über R bzw. $\mathfrak{o}$ mit K multipliziert. Bezeichnet d die Differentiation von K über k_0, so kann also $M(R/k_0)$ mit dem Modul $\langle R\,dR\rangle \subseteqq M(K/k_0)$ identifiziert werden, entsprechend $M(\mathfrak{o}/k_0)$ mit $\langle \mathfrak{o}\,d\,\mathfrak{o}\rangle$. Anders ausgedrückt: Die Differentiation von R über k_0 bzw. von $\mathfrak{o}$ über k_0 ist die Einschränkung der Differentiation von K über k_0 auf R bzw. $\mathfrak{o}$.

Es sei $du_1, \ldots, du_t$ eine Basis von $M(K/k_0)$ über K. $A(K/k_0)$ bezeichne die äußere Algebra von $M(K/k_0)$. Ihre Elemente ω lassen sich eindeutig in der Form

$$\omega = \varkappa_0 + \sum_{\substack{i_1 < \cdots < i_s \\ 1 \leqq s \leqq t}} \varkappa_{i_1 \ldots i_s}\, du_{i_1} \wedge \cdots \wedge du_{i_s} \qquad (\varkappa_0, \varkappa_{i_1 \ldots i_s} \in K) \tag{1}$$

schreiben. Wir nennen ω eine *Differentialform von K über k_0*. Ist ω in $du_1, \ldots, du_t$ homogen vom Grad q, dann heißt ω eine Differentialform q-ter Stufe. Die Elemente von K sind die Differentialformen 0-ter Stufe, die Elemente des Differentialmoduls $M(K/k_0)$ die Differentialformen 1. Stufe. Unter den *Differentialformen höchster Stufe* verstehen wir die t-ter Stufe, die alle von der Form $\varkappa \cdot du_1 \wedge \cdots \wedge du_t$ $(\varkappa \in K)$ sind. Für $q > t$ ist jede Differentialform q-ter Stufe gleich 0.

Wir wollen die Differentialformen auf V untersuchen. Zu diesem Zweck ordnen wir jeder Differentialform ω zunächst für jeden Primdivisor 1. Art auf V einen Wert zu.

DEFINITION. Es sei R der Bewertungsring eines Primdivisors 1. Art auf V. Elemente $\pi_1, \ldots, \pi_t$ aus R heißen ein *System uniformisierender Parameter von R über k_0*, wenn $d\pi_1, \ldots, d\pi_t$ eine Basis von $M(R/k_0)$ bilden.

Ein solches System uniformisierender Parameter gibt es offenbar immer. $d\pi_1, \ldots, d\pi_t$ bilden dann auch eine Basis von $M(K/k_0)$ über K, weshalb jedes Element $\omega \in A(K/k_0)$ auch eine Darstellung (1) mit π anstelle von u besitzt.

DEFINITION. $\omega \in A(K/k_0)$ sei mit Hilfe der uniformisierenden Parameter $\pi_1, \ldots, \pi_t$ für R über k_0 in der Form (1) dargestellt. v sei die zu R gehörige Bewertung. Dann ist

$$v(\omega) = \operatorname*{Min}_{\substack{i_1 < \cdots < i_s \\ 1 \leqq s \leqq t}} \{v(\varkappa_0), v(\varkappa_{i_1 \ldots i_s})\}.$$

Es ist zu zeigen, daß diese Definition unabhängig ist von der Wahl der uniformisierenden Parameter von R über k_0. Es sei $\pi_1', \ldots, \pi_t'$ ein weiteres solches System. Dann gilt

$$\{d\pi_1', \ldots, d\pi_t'\} = \{d\pi_1, \ldots, d\pi_t\} \cdot P$$

mit einer t-reihigen unimodularen Matrix P mit Koeffizienten aus R. Es sei

$$\omega = \varkappa_0' + \sum_{\substack{i_1 < \cdots < i_s \\ 1 \leqq s \leqq t}} \varkappa_{i_1 \ldots i_s}'\, d\pi_{i_1}' \wedge \cdots \wedge d\pi_{i_s}' \qquad (\varkappa_0', \varkappa_{i_1 \ldots i_s}' \in K).$$

Man sieht sofort, daß sich $\varkappa_0$ und die $\varkappa_{i_1 \dots i_s}$ linear durch $\varkappa_0'$ und die $\varkappa'_{i_1 \dots i_s}$ mit Koeffizienten aus R ausdrücken lassen. Mithin ist

$$\operatorname*{Min}_{\substack{i_1 < \dots < i_s \\ 1 \leqq s \leqq t}} \{v(\varkappa_0), v(\varkappa_{i_1 \dots i_s})\} \geqq \operatorname*{Min}_{\substack{i_1 < \dots < i_s \\ 1 \leqq s \leqq t}} \{v(\varkappa_0'), v(\varkappa'_{i_1 \dots i_s})\}.$$

Aus Symmetriegründen gilt auch die umgekehrte Ungleichung. Damit ist bewiesen, daß die Definition sinnvoll ist.

SATZ 3. *v durchlaufe alle Bewertungen, die zu Primdivisoren 1. Art auf V gehören. $\omega \neq 0$ sei eine Differentialform von K über k_0. Dann gilt $v(\omega) = 0$ für fast alle v.*

BEWEIS. Wir betrachten zunächst alle Primdivisoren 1. Art auf V, deren Bewertungsringe lokale Ringe eines minimalen Primideals aus einem festen A_i' sind. A_i' ist ein endlicher Erweiterungsring von k_0: $A_i' = k_0[y_1, \dots, y_m]$. Für die genannten Bewertungsringe R gilt daher: $M(R/k_0)$ wird von $dy_1, \dots, dy_m$ erzeugt. Nun kann man aus einem Erzeugendensystem von $M(R/k_0)$ immer eine Basis von $M(R/k_0)$ auswählen, also aus $y_1, \dots, y_m$ ein System uniformisierender Parameter von R über k_0. Dazu gibt es nur endlich viele Möglichkeiten. Um die Werte von ω für die Primdivisoren 1. Art auf V zu bestimmen, genügt es also, endlich viele Darstellungen (1) von ω zu betrachten. Für Elemente $\varkappa$ aus K ist bekannt, daß $v(\varkappa) = 0$ ist für fast alle v. Daraus folgt nun die Behauptung von Satz 3.

Wegen Satz 3 kann jeder Differentialform $\omega \in A(K/k_0)$ ein Divisor (ω) zugeordnet werden: Durchläuft $\mathfrak{p}$ alle Primdivisoren 1. Art auf V, $v_\mathfrak{p}$ die zugehörigen Bewertungen von K über k, so ist $(\omega) = \prod_\mathfrak{p} \mathfrak{p}^{v_\mathfrak{p}(\omega)}$.

DEFINITION. Eine Differentialform ω von K über k_0 heißt von *1. Gattung auf V*, wenn (ω) ein ganzer Divisor ist, d.h. $v(\omega) \geqq 0$ für alle v.

Gleichbedeutend damit ist: Für jeden Bewertungsring R eines Primdivisors 1. Art auf V ist ω ein Element der äußeren Algebra $A(R/k_0)$ von $M(R/k_0)$.

Ist R ein beliebiger Unterring von K, so bezeichnen wir mit $A^*(R/k_0)$ den R-Untermodul aus $A(K/k_0)$, der von 1 und allen Differentialformen $dr_{i_1} \wedge \dots \wedge dr_{i_s}$ mit $r_{i_1}, \dots, r_{i_s} \in R$ $(i_1 < \dots < i_s,\ 1 \leqq s \leqq t)$ erzeugt wird. Ist R Bewertungsring eines Primdivisors 1. Art auf V oder ein regulärer lokaler Ring von V, so ist $A^*(R/k_0) = A(R/k_0)$.

DEFINITION. Ist ω eine Differentialform von K über k_0 und R ein Bewertungsring von K über k oder $R = \mathfrak{o} \in V$, so heißt ω *regulär für R*, wenn $\omega \in A^*(R/k_0)$ ist. ω heißt *reguläre Differentialform von K/k_0*, wenn ω regulär ist für jeden Bewertungsring R von K/k. ω heißt *regulär auf V*, wenn ω für jedes $\mathfrak{o} \in V$ regulär ist.

Ist ω regulär auf V, so ist ω auch eine reguläre Differentialform von K/k_0, weil es zu jedem Bewertungsring R von K/k ein $\mathfrak{o} \in V$ mit $\mathfrak{o} \subseteqq R$ gibt. Eine reguläre Differentialform von K/k_0 ist auch von 1. Gattung auf V. Daher ist „regulär auf V" der stärkste, „von 1. Gattung auf V" der schwächste unter den drei Begriffen.

DEFINITION. Elemente $\pi_1, \ldots, \pi_t \in \mathfrak{o} \in V$ heißen ein *System uniformisierender Parameter für* $\mathfrak{o}$ bezüglich k_0, wenn $\langle \mathfrak{o}\, d\mathfrak{o} \rangle$ ein freier $\mathfrak{o}$-Modul mit der Basis $d\pi_1, \ldots, d\pi_t$ ist.

SATZ 4. *Ist* $\mathfrak{o}$ *ein regulärer Ring von* V, *dann besitzt er ein System uniformisierender Parameter bezüglich* k_0.

Denn in diesem Fall ist $\langle \mathfrak{o}\, d\mathfrak{o} \rangle = M(\mathfrak{o}/k_0)$ ein freier Modul, der dann auch eine aus Differentialen bestehende Basis besitzt.

SATZ 5. *V sei singularitätenfrei (d.h. jedes $\mathfrak{o} \in V$ sei regulär). Dann ist jede Differentialform 1. Gattung auf V auch regulär auf V.*

BEWEIS. Da jedes $\mathfrak{o} \in V$ regulär ist, ist es insbesondere ganz abgeschlossen. Die Bewertungsringe R von Primdivisoren 1. Art auf V, die $\mathfrak{o}$ umfassen, entsprechen daher umkehrbar eindeutig den minimalen Primidealen von $\mathfrak{o}$ und sind die Quotientenringe von $\mathfrak{o}$ nach diesen Idealen. Ferner gilt $\mathfrak{o} = \bigcap_{\mathfrak{o} \subseteq R} R$. $\pi_1, \ldots, \pi_t$ sei ein System uniformisierender Parameter für $\mathfrak{o}$ bezüglich k_0. Dann ist es auch ein System uniformisierender Parameter für jedes R mit $\mathfrak{o} \subseteq R$, da R ein Quotientenring von $\mathfrak{o}$ ist. Es sei

$$\omega = \varrho_0 + \sum_{\substack{i_1 < \cdots < i_s \\ 1 \leq s \leq t}} \varrho_{i_1 \ldots i_s}\, d\pi_{i_1} \wedge \cdots \wedge d\pi_{i_s} \quad (\varrho_0, \varrho_{i_1 \ldots i_s} \in K)$$

von 1. Gattung auf V. Dann ist $\varrho_0, \varrho_{i_1 \ldots i_s} \in R$ für alle R, mithin auch ϱ_0, $\varrho_{i_1 \ldots i_s} \in \mathfrak{o} = \bigcap_{\mathfrak{o} \subseteq R} R$, womit ω als regulär auf V erkannt ist.

FOLGERUNGEN. 1. *Besitzt der Funktionenkörper K/k ein singularitätenfreies vollständiges Modell und ist k_0 ein Differentialkonstantenkörper von V, so sind die Begriffe „reguläre Differentialform auf V über k_0", „reguläre Differentialform von K/k_0", und „Differentialform 1. Gattung auf V" äquivalent.*

2. *Ist V ein singularitätenfreies, V' ein beliebiges vollständiges Modell von K/k und k_0 ein gemeinsamer Differentialkonstantenkörper von V und V', so ist eine Differentialform von K/k_0, die von 1. Gattung auf V ist, auch von 1. Gattung auf V'.*

Es sei V ein beliebiges vollständiges Modell von K/k und k_0 ein Differentialkonstantenkörper von V. Wir betrachten die Differentialformen höchster Stufe von K/k_0. Da sie einen zyklischen K-Modul bilden, bestehen ihre Divisoren auf V aus allen Divisoren einer Divisorenklasse. Wir zeigen

SATZ 6. *Die Klasse der Divisoren auf V der Differentialformen höchster Stufe von K über k_0 hängt nicht vom Differentialkonstantenkörper k_0 ab.*

BEWEIS. Es seien k_0 und k_1 zwei Differentialkonstantenkörper für V. Nach Hilfssatz 4 ist $[k : k_0]$ und $[k : k_1]$ endlich. $\alpha_1, \ldots, \alpha_r$ sei eine p-Basis von k über k_0, $\beta_1, \ldots, \beta_s$ eine p-Basis von k über k_1. V sei die Menge aller lokalen Ringe der Primideale der affinen k-Algebren $A_1, \ldots, A_m$, deren ganze Abschließungen mit $\bar{A}_1, \ldots, \bar{A}_m$ bezeichnet seien. Dann gibt es einen gemeinsamen Definitionskörper k' von $A_1, \ldots, A_m, \bar{A}_1, \ldots, \bar{A}_m$, der überdies $\alpha_1, \ldots, \alpha_r$, $\beta_1, \ldots, \beta_s$ enthält. Es folgt $k' k_0 = k' k_1 = k$. Aus jeder p-Basis von k_0 bzw. k_1 über k^p läßt sich daher eine p-Basis $\{x^{(0)}\}$ bzw. $\{x^{(1)}\}$ von k über k' auswählen.

Nach Hilfssatz 2 sind dann die Körper $k_0' = k^p[\{x^{(0)}\}]$ und $k_1' = k^p[\{x^{(1)}\}]$ selbst Differentialkonstantenkörper von V und es ist $k_0' \subseteq k_0$ und $k_1' \subseteq k_1$. Es genügt daher, den Beweis für die folgenden beiden Fälle zu führen:

a) k_0 ist in k_1 enthalten,

b) k_0 und k_1 sind zwei Differentialkonstantenkörper der Art, wie sie in Hilfssatz 2 auftreten, die zum gleichen Definitionskörper k' gehören.

a) Es sei $\eta_1, \ldots, \eta_s$ eine p-Basis von k_1 über k_0 und R der Bewertungsring eines Primdivisors 1. Art auf V. d sei die Differentiation von K, d_0 die von K über k_0, d_1 die von K über k_1. Aus der Definition des Differentialkonstantenkörpers und Satz 1 folgt dann

$$M(R) = \langle R\,dR\rangle = E \oplus \overset{s}{\underset{i=1}{\oplus}} R\,d\eta_i \oplus \langle R\,dk_0\rangle,$$

wo E ein endlicher, freier R-Modul ist, der isomorph zu $M(R/k_1)$ ist, während $M(R/k_0) \cong E \oplus \overset{s}{\underset{i=1}{\oplus}} R\,d\eta_i$ ist. Bildet $d_1\pi_1, \ldots, d_1\pi_t$ eine Basis von $M(R/k_1)$, dann ist $d_0\pi_1, \ldots, d_0\pi_t, d_0\eta_1, \ldots, d_0\eta_s$ eine Basis von $M(R/k_0)$. Ist $d_1x_1, \ldots, d_1x_t$ eine Basis von $M(K/k_1)$, so ist $d_0x_1, \ldots, d_0x_t, d_0\eta_1, \ldots, d_0\eta_s$ eine Basis von $M(K/k_0)$. Die Determinante der Übergangsmatrix von der Basis $\{d_0x_1, \ldots, d_0x_t, d_0\eta_1, \ldots, d_0\eta_s\}$ zur Basis $\{d_0\pi_1, \ldots, d_0\pi_t, d_0\eta_1, \ldots, d_0\eta_s\}$ von $M(K/k_0)$ stimmt mit der Determinante der Übergangsmatrix von $\{d_1x_1, \ldots, d_1x_t\}$ zu $\{d_1\pi_1, \ldots, d_1\pi_t\}$ überein. Hieraus ergibt sich, daß die Divisoren der Differentialformen $d_0x_1 \wedge \cdots \wedge d_0x_t \wedge d_0\eta_1 \wedge \cdots \wedge d_0\eta_s$ und $d_1x_1 \wedge \cdots \wedge d_1x_t$ übereinstimmen.

b) Hier ergibt sich aus Hilfssatz 2 sowie Satz 1

$$M(R) = \langle R\,dR\rangle = E_R \oplus \langle R\,dk_0\rangle = E_R \oplus \langle R\,dk_1\rangle$$

mit einem endlichen, freien R-Modul E_R, der nach Hilfssatz 2 eine aus Differentialen bestehende Basis besitzt. Wir zeigen zunächst: Die Moduln E_R aller Bewertungsringe R der Primdivisoren 1. Art auf V erzeugen in $M(K)$ alle denselben K-Modul $K \cdot E_R = E$. Dies ist auf Grund von Hilfssatz 2 klar für zwei Bewertungsringe, die Quotientenringe derselben affinen k-Algebra sind. Es seien nun R und R' zwei beliebige Bewertungsringe von Primdivisoren 1. Art auf V und $\mathfrak{o}$ bzw. $\mathfrak{o}'$ ihre Zentren auf V. Dann liegt $\mathfrak{o}$ in einem affinen Modell V_a von V und $\mathfrak{o}'$ in einem affinen Modell V_a' von V. Der Durchschnitt $V_a \cap V_a'$ ist nicht leer (weil $K \in V_a, V_a'$) und daher selbst ein abstraktes Modell von K/k (vgl. [2], II, Satz 1). Es gibt daher einen Primdivisor 1. Art auf V, dessen Zentrum auf $V_a \cap V_a'$ liegt. Sein Bewertungsring sei R''. Dann sind R und R'' sowie R' und R'' jeweils Quotientenringe einer affinen k-Algebra. Es folgt $K \cdot E_R = K \cdot E_{R''}$ und $K \cdot E_{R'} = K \cdot E_{R''}$, also $K \cdot E_R = K \cdot E_{R'}$ q.e.d.

d_0 sei die Differentiation von K über k_0, d_1 die von K über k_1. Ist $dx_1, \ldots, dx_t$ eine Basis von E, dann ist $d_ix_1, \ldots, d_ix_t$ eine Basis von $M(K/k_i)$ $(i = 0, 1)$. Ist $d\pi_1, \ldots, d\pi_t$ eine Basis von E_R, so ist $d\pi_1, \ldots, d_i\pi_t$ eine Basis von $M(R/k_i)$ $(i = 0, 1)$. Aus $dx_1 \wedge \cdots \wedge dx_t = \varkappa \cdot d\pi_1 \wedge \cdots \wedge d\pi_t$ folgt $d_ix_1 \wedge \cdots \wedge d_ix_t = \varkappa \cdot d_i\pi_1 \wedge \cdots \wedge d_i\pi_t$ $(i = 0, 1)$. Daher stimmen die Divisoren von $d_0x_1 \wedge \cdots \wedge d_0x_t$ und $d_1x_1 \wedge \cdots \wedge d_1x_t$ überein.

DEFINITION. Die nach Satz 6 von der Wahl des Differentialkonstantenkörpers k_0 von V unabhängige Klasse der Divisoren der Differentialformen höchster Stufe von K über k_0 heißt *die kanonische Klasse von V*.

Es sei $\bar{k}$ die algebraische Abschließung von k in K. Die Differentialformen höchster Stufe von K über k_0, die von 1. Gattung auf V sind, bilden einen $\bar{k}$-Modul, der isomorph ist zum $\bar{k}$-Modul aller Vielfachen eines Divisors aus der kanonischen Klasse. Von letzterem ist bekannt, daß seine Dimension über $\bar{k}$ endlich ist.

DEFINITION. Die Dimension des $\bar{k}$-Moduls der Differentialformen höchster Stufe von K über k_0, die von 1. Gattung auf V sind, heißt das *geometrische Geschlecht von V*. Es wird mit $p_g(V)$ bezeichnet.

Nach Satz 6 hängt $p_g(V)$ nicht vom Differentialkonstantenkörper k_0 von V ab. Ist V ein singularitätenfreies, V' ein beliebiges vollständiges Modell von K/k so folgt $p_g(V) \leqq p_g(V')$ nach Satz 5, Folgerung 2. Zwei singularitätenfreie vollständige Modelle von K/k haben also immer das gleiche geometrische Geschlecht.

DEFINITION. Unter dem *geometrischen Geschlecht $p_g(K/k)$ von K über k* werde das Minimum aller $p_g(V)$ verstanden, wo V alle vollständigen abstrakten Modelle von K/k durchläuft.

Existiert ein singularitätenfreies vollständiges Modell V von K/k, so ist $p_g(K/k) = p_g(V)$. Es wird sich später zeigen, daß das hier eingeführte geometrische Geschlecht $p_g(K/k)$ für einen Funktionenkörper K einer Variablen über k mit dem im Riemann-Rochschen Satz auftretenden Geschlecht von K übereinstimmt.

§ 4. Verallgemeinerungen der Riemannschen Formel

Ist K ein separabel erzeugbarer algebraischer Funktionenkörper einer Variablen über k und x ein separierendes transzendentes Element von K/k, dann gilt für den Divisor des exakten Differentials dx bekanntlich $(dx) = \mathfrak{d}\left(\frac{K}{k(x)}\right) \cdot \mathfrak{n}_x^{-2}$, wo $\mathfrak{d}$ die (Dedekindsche) Differente bedeutet und $\mathfrak{n}_x$ den Nennerdivisor von x (vgl. [3], S. 110). Im folgenden Satz wird eine Verallgemeinerung dieser nach RIEMANN benannten Formel gegeben.

SATZ 7. *Es sei K ein algebraischer Funktionenkörper über k und V ein vollständiges abstraktes Modell von K/k. k_0 sei ein Differentialkonstantenkörper von V und $dx_1, \ldots, dx_t$ eine Basis von $M(K/k_0)$. $\mathfrak{p}$ durchlaufe alle Primdivisoren 1. Art auf V, $R_\mathfrak{p}$ die zugehörigen Bewertungsringe. Für jedes $\mathfrak{p}$ sei $\varepsilon_{i\mathfrak{p}} = \pm 1$ so bestimmt, daß $x_i^{\varepsilon_{i\mathfrak{p}}} \in R_\mathfrak{p}$ ist $(i = 1, \ldots, t)$. Ferner sei $\mathfrak{n}_{x_i}$ der Nennerdivisor von x_i $(i = 1, \ldots, t)$ und $\mathfrak{d}_j$ bezeichne die j-te Kählersche Differente. Dann gilt für den Divisor der Differentialform $dx_{i_1} \wedge \cdots \wedge dx_{i_s}$ $(i_1 < i_2 < \cdots < i_s,\ 1 \leqq s \leqq t)$ die Formel*

$$(dx_{i_1} \wedge \cdots \wedge dx_{i_s}) = \frac{\prod_{\mathfrak{p}} \mathfrak{d}_{t-s}\left(\frac{R_\mathfrak{p}}{k_0\left[x_{i_1}^{\varepsilon_{i_1\mathfrak{p}}}, \ldots, x_{i_s}^{\varepsilon_{i_s\mathfrak{p}}}\right]}\right)}{\mathfrak{n}_{x_{i_1}}^2 \cdot \cdots \cdot \mathfrak{n}_{x_{i_s}}^2}$$

BEWEIS. $\pi_1, \ldots, \pi_t$ sei ein System uniformisierender Parameter von $R_\mathfrak{p}$ bezüglich k_0. Wegen $dx_1^{e_1\mathfrak{p}}, \ldots, dx_t^{e_t\mathfrak{p}} \in dR_\mathfrak{p}$ gilt dann

$$\{dx_1^{e_1\mathfrak{p}}, \ldots, dx_t^{e_t\mathfrak{p}}\} = \{d\pi_1, \ldots, d\pi_t\} \cdot P$$

mit einer t-reihigen quadratischen Matrix P mit Koeffizienten aus $R_\mathfrak{p}$. Drückt man nun $dx_{i_1}^{e_{i_1}\mathfrak{p}} \wedge \cdots \wedge dx_{i_s}^{e_{i_s}\mathfrak{p}}$ durch die $d\pi_{j_1} \wedge \cdots \wedge d\pi_{j_s}$ aus, so erhält man als Koeffizienten bis aufs Vorzeichen gerade alle s-reihigen Unterdeterminanten, die sich aus den i_1-ten, ..., i_s-ten Spalten der Matrix P bilden lassen. Es ist dann $\nu_\mathfrak{p}(dx_{i_1}^{e_{i_1}\mathfrak{p}} \wedge \cdots \wedge dx_{i_s}^{e_{i_s}\mathfrak{p}})$ gleich dem Minimum der Werte dieser Unterdeterminanten.

Andererseits gilt $M\left(\frac{R_\mathfrak{p}}{k_0[x_{i_1}^{e_{i_1}\mathfrak{p}}, \ldots, x_{i_s}^{e_{i_s}\mathfrak{p}}]}\right) \cong M\left(\frac{R_\mathfrak{p}}{k_0}\right) / \langle dx_{i_1}^{e_{i_1}\mathfrak{p}}, \ldots, dx_{i_s}^{e_{i_s}\mathfrak{p}} \rangle$ und $M\left(\frac{R_\mathfrak{p}}{k_0}\right)$ ist ein freier $R_\mathfrak{p}$-Modul mit der Basis $d\pi_1, \ldots, d\pi_t$. Die Kählerschen Differenten von $R_\mathfrak{p}$ über $k_0[x_{i_1}^{e_{i_1}\mathfrak{p}}, \ldots, x_{i_s}^{e_{i_s}\mathfrak{p}}]$ berechnen sich daher nach ihrer Definition aus den Unterdeterminanten der Koeffizientenmatrix, mit der man $dx_{i_1}^{e_{i_1}\mathfrak{p}}, \ldots, dx_{i_s}^{e_{i_s}\mathfrak{p}}$ durch $d\pi_1, \ldots, d\pi_t$ ausdrückt. Die s-reihigen Unterdeterminanten liefern die Differente $\mathfrak{d}_{t-s}$. Damit ist gezeigt, daß die $\mathfrak{p}$-Komponente des Divisors von $dx_{i_1}^{e_{i_1}\mathfrak{p}} \wedge \cdots \wedge dx_{i_s}^{e_{i_s}\mathfrak{p}}$ mit $\mathfrak{d}_{t-s}\left(\frac{R_\mathfrak{p}}{k_0[x_{i_1}^{e_{i_1}\mathfrak{p}}, \ldots, x_{i_s}^{e_{i_s}\mathfrak{p}}]}\right)$ übereinstimmt.

Wegen $dx^{-1} = -x^{-2}dx$ folgt nun sofort

$$dx_{i_1} \wedge \cdots \wedge dx_{i_s} = \pm x_{i_1}^{1-e_{i_1}\mathfrak{p}} \cdot \cdots \cdot x_{i_s}^{1-e_{i_s}\mathfrak{p}} \cdot dx_{i_1}^{e_{i_1}\mathfrak{p}} \wedge \cdots \wedge dx_{i_s}^{e_{i_s}\mathfrak{p}}.$$

Hieraus ergibt sich unmittelbar die Behauptung von Satz 7.

Bemerkung. Satz 7 gilt natürlich auch für Char. $k = 0$ mit $k_0 = k$.

Man kann auch den Differentialformen von K über dem arithmetischen Konstantenkörper k in analoger Weise wie den Differentialformen von K über einem Differentialkonstantenkörper k_0 für jeden Primdivisor 1. Art auf V einen Wert zuordnen. Ist $dx_1, \ldots, dx_t$ eine Basis von $M(K/k)$, so wird $-\nu_\mathfrak{p}(dx_1 \wedge \cdots \wedge dx_t)$ definiert als das Minimum aller ν mit $\mathfrak{p}^\nu \cdot (dx_1 \wedge \cdots \wedge dx_t) \subseteq R_\mathfrak{p} dR_\mathfrak{p} \wedge \cdots \wedge dR_\mathfrak{p}$. Für den hierdurch gegebenen Divisor der Differentialform $dx_1 \wedge \cdots \wedge dx_t$ bewies F. K. SCHMIDT die Formel (unveröffentlicht)

$$(dx_1 \wedge \cdots \wedge dx_t) = \frac{\prod_\mathfrak{p} \mathfrak{d}_0\left(\frac{R_\mathfrak{p}}{k[x_1^{e_1\mathfrak{p}}, \ldots, x_t^{e_t\mathfrak{p}}]}\right)}{\prod_\mathfrak{p} \mathfrak{d}_t\left(\frac{R_\mathfrak{p}}{k}\right) \cdot \mathfrak{n}_{x_1}^2 \cdot \cdots \cdot \mathfrak{n}_{x_t}^2},$$

wobei die Bezeichnungen sinngemäß wie in Satz 7 zu wählen sind. Die durch diese Differentialformen gegebene „kanonische Klasse" erweist sich jedoch im Fall algebraischer Funktionenkörper einer Variablen als nicht identisch mit der kanonischen Klasse des Riemann-Rochschen Satzes.

F. K. SCHMIDT weist mich noch darauf hin, daß die Differente $\left(\frac{R_\mathfrak{p}}{k_0[x_1^{e_1\mathfrak{p}}, \ldots, x_t^{e_t\mathfrak{p}}]}\right)$ eine Schachtelung der folgenden Form gestattet:

$$\mathfrak{d}_0\left(\frac{R_\mathfrak{p}}{k_0[x_1^{e_1\mathfrak{p}}, \ldots, x_t^{e_t\mathfrak{p}}]}\right) = \mathfrak{d}_0\left(\frac{R_\mathfrak{p}}{P_\mathfrak{p}}\right) \cdot \mathfrak{d}_0\left(\frac{P_\mathfrak{p}}{k_0[x_1^{e_1\mathfrak{p}}, \ldots, x_t^{e_t\mathfrak{p}}]}\right).$$

Hierbei ist $P_{\mathfrak{p}}$ der zu $\mathfrak{p}$ gehörige Bewertungsring von $k_0(x_1, \ldots, x_t)$. Der Beweis für diese Tatsache ergibt sich aus der Fortsetzungstheorie der Derivationen. Die Formel liefert eine genauere Einteilung der Primdivisoren, die in $(dx_1 \wedge \cdots \wedge dx_t)$ aufgehen, als es nach Satz 7 möglich ist.

§ 5. Änderung der kanonischen Klasse bei endlich algebraischer Erweiterung des Funktionenkörpers

Es sei K/k ein algebraischer Funktionenkörper, V ein vollständiges Modell von K/k. L sei ein endlicher algebraischer Oberkörper von K und V' das abgeleitete normale Modell von V in L, d.h. ist V die Menge aller lokalen Ringe von Primidealen der affinen k-Algebren $A_1, \ldots, A_s$, so ist V' die Menge aller lokalen Ringe von Primidealen der ganzen Abschließungen $A_1', \ldots, A_s'$ der A_i in L (vgl. [1]). Das Ziel dieses Paragraphen ist die Herleitung eines Zusammenhangs der kanonischen Klassen von V und V'.

Ist $\mathfrak{p}$ ein Primdivisor 1. Art auf V, so erhält man durch die Fortsetzungen seiner Bewertung auf L Primdivisoren $\mathfrak{P}_1, \ldots, \mathfrak{P}_k$ 1. Art auf V'. Auf diese Weise bekommt man auch alle Primdivisoren 1. Art auf V'. Sind $e_1, \ldots, e_k$ die Verzweigungsexponenten der $\mathfrak{P}_i$ über $\mathfrak{p}$, so wird $\mathfrak{p}$ auf V' der Erweiterungsdivisor $\mathfrak{P}_1^{e_1} \cdot \cdots \cdot \mathfrak{P}_k^{e_k}$ zugeordnet. Es ist klar, was unter dem Erweiterungsdivisor $\mathfrak{A}$ auf V' eines beliebigen Divisors $\mathfrak{a}$ von V zu verstehen ist.

Wir führen die Untersuchung in zwei Schritten durch, indem wir zuerst separabel algebraische Erweiterungen studieren, dann rein inseparable Erweiterungen.

a) Separable algebraische Erweiterungen

Es sei k_0 ein gemeinsamer Differentialkonstantenkörper von V und V'. Ferner sei $M(K/k_0) = K\,dx_1 \oplus \cdots \oplus K\,dx_t$. Ist L separabel algebraisch über K, dann ist $M(L/k_0) = L\,dx_1 \oplus \cdots \oplus L\,dx_t$, wobei die Differentiation $d: L \to M(L/k_0)$ Fortsetzung der Differentiation $d: K \to M(K/k_0)$ ist. Um die Änderung der kanonischen Klasse zu überblicken, genügt es daher, den Divisor von $dx_1 \wedge \cdots \wedge dx_t$ auf V mit dem von $dx_1 \wedge \cdots \wedge dx_t$ auf V' zu vergleichen. Es sei $S_{\mathfrak{P}}$ der Bewertungsring eines Primdivisors $\mathfrak{P}$ 1. Art auf V' und $R_{\mathfrak{p}} = S_{\mathfrak{P}} \cap K$. $\pi_1, \ldots, \pi_t$ sei ein System uniformisierender Parameter von $S_{\mathfrak{P}}$ über k_0, $\varrho_1, \ldots, \varrho_t$ ein System uniformisierender Parameter von $R_{\mathfrak{p}}$ über k_0 und es gelte $d\varrho_1 \wedge \cdots \wedge d\varrho_t = s \cdot d\pi_1 \wedge \cdots \wedge d\pi_t$. Dann ist s die Determinante der Matrix mit Koeffizienten aus $S_{\mathfrak{P}}$, mit der man $d\varrho_1, \ldots, d\varrho_t$ durch $d\pi_1, \ldots, d\pi_t$ darstellt. Wegen $M(S_{\mathfrak{P}}/R_{\mathfrak{p}}) \cong M(S_{\mathfrak{P}}/k_0)/\langle S_{\mathfrak{P}} dR_{\mathfrak{p}} \rangle = M(S_{\mathfrak{P}}/k_0)/\langle d\varrho_1, \ldots, d\varrho_t \rangle$ erzeugt s auch die 0-te Kählersche Differente $\mathfrak{d}_0(S_{\mathfrak{P}}/R_{\mathfrak{p}})$. Für die $\mathfrak{P}$-Komponenten der Divisoren von $d\varrho_1 \wedge \cdots \wedge d\varrho_t$ und $d\pi_1 \wedge \cdots \wedge d\pi_t$ auf V' ergibt sich

$$(d\varrho_1 \wedge \cdots \wedge d\varrho_t)_{\mathfrak{P}} = \mathfrak{d}_0\left(\frac{S_{\mathfrak{P}}}{R_{\mathfrak{p}}}\right)(d\pi_1 \wedge \cdots \wedge d\pi_t)_{\mathfrak{P}}.$$

Aus

$$dx_1 \wedge \cdots \wedge dx_t = \varkappa \cdot d\varrho_1 \wedge \cdots \wedge d\varrho_t = \varkappa \cdot s \cdot d\pi_1 \wedge \cdots \wedge d\pi_t$$

folgt

$$v_{\mathfrak{P}}(dx_1 \wedge \cdots \wedge dx_t) = v_{\mathfrak{P}}(\varkappa) + v_{\mathfrak{P}}\left(\mathfrak{d}_0\left(\frac{S_{\mathfrak{P}}}{R_{\mathfrak{p}}}\right)\right).$$

Hieraus folgt nun sofort

SATZ 8. *L sei separabel algebraisch über K. Ist* $\mathfrak{k}$ *ein Divisor aus der kanonischen Klasse von V,* $\mathfrak{K}$ *sein Erweiterungsdivisor auf V', so ist* $\mathfrak{K} \cdot \prod_{\mathfrak{P}} \mathfrak{d}_0(S_{\mathfrak{P}}/R_{\mathfrak{p}})$ *ein Divisor der kanonischen Klasse von V'. Dabei ist das Produkt über alle Primdivisoren 1. Art auf V' zu erstrecken.*

Wir schreiben $\prod_{\mathfrak{P}} \mathfrak{d}_0(S_{\mathfrak{P}}/R_{\mathfrak{p}}) = \mathfrak{d}_0(V'/V)$ und nennen $\mathfrak{d}_0(V'/V)$ die 0-te Kählersche Differente von V' über V.

b) Differentialformen und Tatesche Spur

Bevor wir zu inseparablen Körpererweiterungen übergehen, leiten wir einen Zusammenhang zwischen den Differentialformen und der sog. Tateschen Spur her.

Wegen $M(K/k_0) = K\,dx_1 \oplus \cdots \oplus K\,dx_t$ bilden $x_1, \ldots, x_t$ eine p-Basis von K über $K' = k_0(K^p)$. Wir setzen $K_0 = K'$, $K_1 = K'[x_1], \ldots, K_t = K'[x_1, \ldots, x_t] = K$. R sei der Bewertungsring eines Primdivisors $\mathfrak{p}$ 1. Art auf V, R_i sei der Bewertungsring von $\mathfrak{p}$ in K_i $(i = 0, \ldots, t)$. Wegen $[K_i : K_{i-1}] = p$ ist K_i über K_{i-1} bei $v_{\mathfrak{p}}$ entweder unverzweigt oder reinverzweigt. In jedem Fall folgt, daß R_i über R_{i-1} von einem Element y_i erzeugt wird: $R_i = R_{i-1}[y_i]$ $(i = 1, \ldots, t)$. Da $R = R_0[y_1, \ldots, y_t]$ ist, bilden $y_1, \ldots, y_t$ ein System uniformisierender Parameter von R über k_0.

Es sei $x_i = \varrho_{0i} + \varrho_{1i} y_i + \cdots + \varrho_{p-1,i} y_i^{p-1}$ mit $\varrho_{ji} \in K_{i-1}$ $(i = 1, \ldots, t)$. Dann gilt

$$dx_i = (\varrho_{1i} + 2\varrho_{2i} y_i + \cdots + (p-1)\varrho_{p-1,i} y_i^{p-2})\,dy_i + \\ + d\varrho_{0i} + y_i\,d\varrho_{1i} + \cdots + y_i^{p-1}\,d\varrho_{p-1,i}.$$

Wir setzen

$$\varrho_{1i} + 2\varrho_{2i} y_i + \cdots + (p-1)\varrho_{p-1,i} y_i^{p-2} = D_{y_i} x_i.$$

Da $dy_1, \ldots, dy_{i-1}$ eine Basis von $K_{i-1}\,dK_{i-1}$ bilden, ergibt sich

$$dx_i = (D_{y_i} x_i) \cdot dy_i + \varkappa_{1i}\,dy_1 + \cdots + \varkappa_{i-1,i}\,dy_{i-1} \quad \text{mit} \quad \varkappa_{ji} \in K \quad (i = 1, \ldots, t).$$

Es folgt

$$dx_1 \wedge \cdots \wedge dx_t = \prod_{i=1}^{t} (D_{y_i} x_i) \cdot dy_1 \wedge \cdots \wedge dy_t. \tag{1}$$

Die Elemente $1, x_i, \ldots, x_i^{p-1}$ bilden eine Basis von K_i über K_{i-1}. Unter der *Tateschen Spur* σ_{x_i} von K_i über K_{i-1} versteht man die lineare Abbildung $\sigma_{x_i}: K_i \to K_{i-1}$, die jedem Element $\varkappa = \varkappa_0 + \varkappa_1 x_i + \cdots + \varkappa_{p-1} x_i^{p-1} \in K_i$ $(\varkappa_0, \ldots, \varkappa_{p-1} \in K_{i-1})$ das Element $\sigma_{x_i}(\varkappa) = \varkappa_{p-1} \in K_{i-1}$ zuordnet. Wir bestimmen die σ_{x_i}-Differente von R_i über R_{i-1}. Dazu haben wir zunächst den Komplementärmodul E_i von R_i zu bestimmen. Da $1, y_i, \ldots, y_i^{p-1}$ eine Basis von R_i über R_{i-1} bilden, findet man eine Basis $z_0, \ldots, z_{p-1}$ von E_i über R_{i-1} durch Auflösung des Gleichungssystems

$$\sigma_{x_i}(z_j \cdot y_i^k) = \delta_{jk}.$$

Wenden wir die Transformationsformel von TATE an (vgl. [6], Satz 1), so erhalten wir

$$\sigma_{x_i}(z_j \cdot y_i^k) = \sigma_{y_i}\left(z_j \cdot y_i^k \cdot (D_{y_i} x_i)^{1-p}\right).$$

Aus der Definition der Tateschen Spur ergibt sich nun sofort, daß

$$z_j = (D_{y_i} x_i)^{p-1} y_i^{p-j-1} \qquad (j = 0, 1, \ldots, p-1)$$

eine Lösung des Gleichungssystems ist. Daher ist $E_i = (D_{y_i} x_i)^{p-1} \cdot R_i$. Für die σ_{x_i}-Differente $\mathfrak{d}_i$ von R_i über R_{i-1} folgt hieraus

$$\mathfrak{d}_i = (D_{y_i} x_i)^{1-p} \cdot R_i. \tag{2}$$

Es sei nun $\sigma: K \to K'$ die lineare Funktion, die durch

$$\sigma(\varkappa) = \sigma_{x_1} \circ \sigma_{x_2} \circ \cdots \circ \sigma_{x_t}(\varkappa) \quad \text{für alle } \varkappa \in K$$

definiert wird. Für die mit Hilfe dieser Linearform gebildete Differente $\mathfrak{d}$ von R über R_0 gilt der Schachtelungssatz (vgl. [5], Satz 1)

$$\mathfrak{d} = \prod_{i=1}^{t} \mathfrak{d}_i.$$

Aus (2) folgt

$$\mathfrak{d} = \prod_{i=1}^{t} (D_{y_i} x_i)^{1-p} \cdot R. \tag{3}$$

(3) und (1) ergeben zusammen

SATZ 9. *$\mathfrak{d}$ sei die mit Hilfe der Linearform $\sigma = \sigma_{x_1} \circ \cdots \circ \sigma_{x_t}$ gebildete Differente von R über R_0. Dann ist $\mathfrak{d} = (d x_1 \wedge \cdots \wedge d x_t)_{\mathfrak{p}}^{1-p}$, wo $(d x_1 \wedge \cdots \wedge d x_t)_{\mathfrak{p}}$ die $\mathfrak{p}$-Komponente des Divisors $(d x_1 \wedge \cdots \wedge d x_t)$ bedeutet.*

c) Inseparable Erweiterungen vom Grad p

Es sei $L = K(x)$ mit $x^p \in K$, $x \notin K$. k' sei ein gemeinsamer Definitionskörper der affinen k-Algebren $A_1, \ldots, A_s$ und $A_1', \ldots, A_s'$, deren lokale Ringe gerade V und V' ausmachen. $\{\varkappa_\mu\}_{\mu \in M}$ sei eine p-Basis von k über k'. Wir wählen als gemeinsamen Differentialkonstantenkörper von V und V' den Körper $k_0 = k^p[\{\varkappa_\mu\}_{\mu \in M}]$. Wegen $L^p \subseteq K$ gilt $K^p[\{\varkappa_\mu\}] = k_0(K^p) \subseteq L^p[\{\varkappa_\mu\}] = k_0(L^p) \subseteq K \subseteq L$. Wir zeigen zunächst, daß die $\varkappa_\mu$ über L^p p-unabhängig sind (dann natürlich auch über K^p). Es sei $A' = k[x_1, \ldots, x_n]$ einer der Ringe A_i'. Nach Hilfssatz 1 sind die $\varkappa_\mu$ ($\mu \in M$) über $k'[x_1, \ldots, x_n]$ p-unabhängig. Dann sind sie aber auch über $L' = k'(x_1, \ldots, x_n)$ p-unabhängig und wegen $L^p = k^p(x_1^p, \ldots, x_t^p) \subseteq L'$ auch über L^p. Es folgt

$$L^p[\{\varkappa_\mu\}] = K^p[\{\varkappa_\mu\}][x^p] \quad \text{mit} \quad x^p \notin K^p[\{\varkappa_\mu\}].$$

HILFSSATZ 6. *Es sei S der Bewertungsring eines Primdivisors 1. Art auf V' und $R = S \cap K$. Dann ist $S \cap L^p[\{\varkappa_\mu\}] = S^p[\{\varkappa_\mu\}]$ und $R \cap K^p[\{\varkappa_\mu\}] = R^p[\{\varkappa_\mu\}]$.*

BEWEIS. Wir zeigen die erste Gleichung, die zweite ergibt sich genau so. S sei Quotientenring von $A' = k[x_1, \ldots, x_n]$ und $S' = S \cap L'$ mit $L' = k'(x_1, \ldots, x_n)$. Wir zeigen zuerst: $S = S'[\{\varkappa_\mu\}]$. Da S^p und k' in S' enthalten sind, gilt

$S^{\mathfrak{p}}[k, x_1, \ldots, x_n] = S^{\mathfrak{p}}[k', x_1, \ldots, x_n][\{x_\mu\}] \subseteq S'[\{x_\mu\}]$. $s \in k[x_1, \ldots, x_n]$ sei ein Element der Nennermenge von S. Dann ist $\frac{1}{s^{\mathfrak{p}}} \in S^{\mathfrak{p}}$. Wegen $s^{\mathfrak{p}-1} \in k[x_1, \ldots, x_n]$ ergibt sich $\frac{1}{s} \in S^{\mathfrak{p}}[k, x_1, \ldots, x_n]$. Es gilt daher $S \subseteq S'[\{x_\mu\}]$ und da die umgekehrte Inklusion trivial ist, folgt $S = S'[\{x_\mu\}]$. Die x_μ $(\mu \in M)$ sind über S' $\mathfrak{p}$-unabhängig, da sie über L' $\mathfrak{p}$-unabhängig sind.

Als nächstes zeigen wir, daß es Elemente $y_1, \ldots, y_t \in S'$ gibt, die über $S^{\mathfrak{p}}$ $\mathfrak{p}$-unabhängig sind, so daß $S' = S^{\mathfrak{p}}[y_1, \ldots, y_t]$ gilt. Es ist nämlich L' eine endliche, rein inseparable Erweiterung des algebraischen Funktionenkörpers $L^{\mathfrak{p}}$, weshalb für S' und $S^{\mathfrak{p}}$ wie in b) für R und R_0 geschlossen werden kann.

Es ergibt sich $S = S^{\mathfrak{p}}[\{x_\mu\}, y_1, \ldots, y_t]$, wobei die Elemente des Systems $\{x_\mu\}, y_1, \ldots, y_t$ $\mathfrak{p}$-unabhängig über $S^{\mathfrak{p}}$ sind und daher auch eine $\mathfrak{p}$-Basis von L über $L^{\mathfrak{p}}$ bilden. Es sei nun $s \in S \cap L^{\mathfrak{p}}[\{x_\mu\}]$. Dann besitzt s als Element von $L^{\mathfrak{p}}[\{x_\mu\}]$ eine Darstellung $s = f(\{x_\mu\})$, wobei f ein Polynom mit Koeffizienten aus $L^{\mathfrak{p}}$ ist, das in allen Variablen vom Grad $\leqq \mathfrak{p} - 1$ ist. Als Element von S besitzt s eine Darstellung $s = g(y_1, \ldots, y_t, \{x_\mu\})$, wo g ein Polynom mit Koeffizienten aus $S^{\mathfrak{p}}$ ist, das in allen Variablen vom Grad $\leqq \mathfrak{p} - 1$ ist. Durch Koeffizientenvergleich folgt hieraus sofort, daß alle Koeffizienten von f in $S^{\mathfrak{p}}$ liegen. Also ist $S \cap L^{\mathfrak{p}}[\{x_\mu\}] \subseteq S^{\mathfrak{p}}[\{x_\mu\}]$. Da die umgekehrte Inklusion trivial ist, folgt die im Hilfssatz behauptete Gleichheit.

Es sei $x_1, \ldots, x_{t-1}$ eine $\mathfrak{p}$-Basis von K über $k_0(L^{\mathfrak{p}}) = L^{\mathfrak{p}}[\{x_\mu\}]$. Dann ist $x, x_1, \ldots, x_{t-1}$ eine $\mathfrak{p}$-Basis von L über $k_0(L^{\mathfrak{p}})$ und $x_1, \ldots, x_{t-1}, x^{\mathfrak{p}}$ eine $\mathfrak{p}$-Basis von K über $k_0(K^{\mathfrak{p}})$. d' sei die Differentiation von L über k_0, d die von K über k_0. Wir vergleichen die Divisoren der Differentialformen $d x_1 \wedge \cdots \wedge d x_{t-1} \wedge d x^{\mathfrak{p}}$ auf V und $d' x \wedge d' x_1 \wedge \cdots \wedge d' x_{t-1}$ auf V'.

Zu diesem Zweck betrachten wir die mit Hilfe der nichttrivialen Linearform $\sigma_{x^{\mathfrak{p}}} \circ \sigma_{x_1} \circ \cdots \circ \sigma_{x_{t-1}} \circ \sigma_x$ gebildete Differente von S über $S \cap K^{\mathfrak{p}}[\{x_\mu\}] = R^{\mathfrak{p}}[\{x_\mu\}]$, wobei die σ jeweils Tatesche Spuren sind. Den beiden Zerlegungen

$$\sigma_{x^{\mathfrak{p}}} \circ (\sigma_{x_1} \circ \cdots \circ \sigma_{x_{t-1}} \circ \sigma_x) = (\sigma_{x^{\mathfrak{p}}} \circ \sigma_{x_1} \circ \cdots \circ \sigma_{x_{t-1}}) \circ \sigma_x$$

entsprechen nach dem Schachtelungssatz Zerlegungen der Differente von folgender Form

$$\mathfrak{d}\left(\frac{S}{S^{\mathfrak{p}}[\{x_\mu\}]}\right) \cdot \mathfrak{d}\left(\frac{S^{\mathfrak{p}}[\{x_\mu\}]}{R^{\mathfrak{p}}[\{x_\mu\}]}\right) = \mathfrak{d}\left(\frac{S}{R}\right) \cdot \mathfrak{d}\left(\frac{R}{R^{\mathfrak{p}}[\{x_\mu\}]}\right). \tag{4}$$

Es sei $\mathfrak{P}$ der Primdivisor von S, $\mathfrak{p}$ der von R. Dann gilt nach Satz 9:

$$\mathfrak{d}\left(\frac{S}{S^{\mathfrak{p}}[\{x_\mu\}]}\right) = (d' x \wedge d' x_1 \wedge \cdots \wedge d' x_{t-1})_{\mathfrak{P}}^{1-\mathfrak{p}},$$

$$\mathfrak{d}\left(\frac{R}{R^{\mathfrak{p}}[\{x_\mu\}]}\right) = (d x_1 \wedge \cdots \wedge d x_{t-1} \wedge d x^{\mathfrak{p}})_{\mathfrak{p}}^{1-\mathfrak{p}}.$$

Wir beweisen jetzt, daß $\mathfrak{d}\left(\frac{S^{\mathfrak{p}}[\{x_\mu\}]}{R^{\mathfrak{p}}[\{x_\mu\}]}\right) \cdot S = \mathfrak{d}\left(\frac{S}{R}\right)^{\mathfrak{p}}$ ist. Wie unter b) gezeigt, gilt $S = R[y] = R \oplus R \cdot y \oplus \cdots \oplus R \cdot y^{\mathfrak{p}-}$ und $\mathfrak{d}(S/R) = (D_y x)^{1-\mathfrak{p}} \cdot S$. Es folgt $S^{\mathfrak{p}}[\{x_\mu\}] = R^{\mathfrak{p}}[\{x_\mu\}][y^{\mathfrak{p}}]$, wo $1, y^{\mathfrak{p}}, \ldots, (y^{\mathfrak{p}})^{\mathfrak{p}-1}$ eine Basis von $S^{\mathfrak{p}}[\{x_\mu\}]$ über

$R^{\mathfrak{p}}[\{x_\mu\}]$ bilden. Folglich ist

$$\mathfrak{d}\left(\frac{S^{\mathfrak{p}}[\{x_\mu\}]}{R^{\mathfrak{p}}[\{x_\mu\}]}\right)\cdot S=(D_{y^{\mathfrak{p}}}x^{\mathfrak{p}})^{1-\mathfrak{p}}\cdot S=[(D_y x)^{1-\mathfrak{p}}]^{\mathfrak{p}}\cdot S=\mathfrak{d}\left(\frac{S}{R}\right)^{\mathfrak{p}}.$$

Aus (4) folgt somit:

$$(d'x\wedge d'x_1\wedge\cdots\wedge d'x_{t-1})_{\mathfrak{P}}^{1-\mathfrak{p}}\,\mathfrak{d}\left(\frac{S}{R}\right)^{\mathfrak{p}}=\mathfrak{d}\left(\frac{S}{R}\right)(dx_1\wedge\cdots\wedge dx_{t-1}\wedge dx^{\mathfrak{p}})_{\mathfrak{P}}^{1-\mathfrak{p}},$$

wo unter $(dx_1\wedge\cdots\wedge dx_{t-1}\wedge dx^{\mathfrak{p}})_{\mathfrak{P}}$ die $\mathfrak{P}$-Komponente des Erweiterungsdivisors von $(dx_1\wedge\cdots\wedge dx_{t-1}\wedge dx^{\mathfrak{p}})$ zu verstehen ist. Es folgt

$$(d'x\wedge d'x_1\wedge\cdots\wedge d'x_{t-1})_{\mathfrak{P}}=\mathfrak{d}\left(\frac{S}{R}\right)\cdot(dx_1\wedge\cdots\wedge dx_{t-1}\wedge dx^{\mathfrak{p}})_{\mathfrak{P}}.$$

Daher unterscheiden sich der Divisor von $d'x\wedge d'x_1\wedge\cdots\wedge d'x_{t-1}$ auf V' und der Erweiterungsdivisor von $dx_1\wedge\cdots\wedge dx_{t-1}\wedge dx^{\mathfrak{p}}$ auf V' nur um $\prod_{\mathfrak{P}}\mathfrak{d}(S_{\mathfrak{P}}/R_{\mathfrak{p}})$, wo $\mathfrak{P}$ alle Primdivisoren 1. Art auf V' durchläuft und $S_{\mathfrak{P}}$ und $R_{\mathfrak{p}}$ die zugehörigen Bewertungsringe in L und K bedeuten. Wir setzen $\prod_{\mathfrak{P}}\mathfrak{d}(S_{\mathfrak{P}}/R_{\mathfrak{p}})=\mathfrak{d}(V'/V)$ und nennen diesen Divisor die σ_x-Differente von V' über V. Bildet man eine σ-Differente $\mathfrak{d}(V'/V)$ bezüglich einer beliebigen nichttrivialen Linearform σ von L über K, so unterscheidet sie sich von der mit Hilfe der Tateschen Spur σ_x gebildeten Differente nur um den Divisor eines Elements aus L (vgl. [5], S. 421). Ist daher $\mathfrak{k}$ ein beliebiger Divisor der kanonischen Klasse von V, $\mathfrak{K}$ sein Erweiterungsdivisor auf V' und $\mathfrak{d}(V'/V)$ die bezüglich einer nichttrivialen Linearform von L über K gebildete Differente von V' über V, so ist $\mathfrak{K}\cdot\mathfrak{d}(V'/V)$ ein Divisor der kanonischen Klasse von V'.

d) Beliebige endlich algebraische Erweiterungen

Im Fall einer endlichen separablen algebraischen Erweiterung L von K ist bekannt, daß die 0-te Kählersche Differente $\mathfrak{d}_0(V'/V)$ mit der Dedekindschen Differente $\mathfrak{d}(V'/V)$ übereinstimmt. Diese ist die mit Hilfe der gewöhnlichen Spur σ gebildete σ-Differente. In Satz 8 darf daher $\mathfrak{d}_0$ durch eine beliebige σ-Differente $\mathfrak{d}$ ersetzt werden. Zusammen mit c) und dem Schachtelungssatz der σ-Differenten folgt dann sofort

SATZ 10. *L sei eine beliebige endliche algebraische Erweiterung von K, $\mathfrak{k}$ ein Divisor der kanonischen Klasse von V, $\mathfrak{K}$ sein Erweiterungsdivisor auf V' und $\mathfrak{d}(V'/V)$ die bezüglich einer nichttrivialen Linearform von L über K gebildete Differente von V' über V. Dann ist $\mathfrak{K}\cdot\mathfrak{d}(V'/V)$ ein Divisor der kanonischen Klasse von V'.*

e) Anwendung auf einen Spezialfall

Es sei $K=k(x_1,\ldots,x_r)$ ein rationaler Funktionenkörper und L eine endliche algebraische Erweiterung von K. V sei das projektive Modell von $k(x_1,\ldots,x_r)$, welches aus allen lokalen Ringen der Primideale der affinen k-Algebren $k[x_1,\ldots,x_r]$, $k[x_1^{-1},x_1^{-1}x_2,\ldots,x_1^{-1}x_r],\ldots,k[x_r^{-1},x_r^{-1}x_1,\ldots,x_r^{-1}x_{r-1}]$ besteht, V' sei das abgeleitete normale Modell von V in L.

Die affinen k-Algebren von V sind ganz abgeschlossen, daher sind die Bewertungsringe der Primdivisoren 1. Art auf V die lokalen Ringe dieser k-Algebren nach ihren minimalen Primidealen. Wir zeigen

HILFSSATZ 7. *Außer den lokalen Ringen der minimalen Primideale von $k[x_1, \ldots, x_r]$ gibt es nur noch einen Bewertungsring, der zu einem Primdivisor 1. Art auf V gehört, nämlich den lokalen Ring des Primideals (x_1^{-1}) in $k[x_1^{-1}, x_1^{-1}x_2, \ldots, x_1^{-1}x_r]$. Ist $\mathfrak{n}_{x_i}$ der Nennerdivisor von x_i auf V $(i=1, \ldots, r)$, dann ist $\mathfrak{n}_{x_1} = \cdots = \mathfrak{n}_{x_r}$.*

BEWEIS. Es sei $\mathfrak{p}$ ein Primdivisor 1. Art auf V mit $k[x_i^{-1}, x_i^{-1}x_1, \ldots, x_i^{-1}x_r] < R_\mathfrak{p}$, dem Bewertungsring von $\mathfrak{p}$. Ist dann $\nu_\mathfrak{p}(x_i^{-1}) = 0$, so folgt sofort $\nu_\mathfrak{p}(x_1), \ldots, \nu_\mathfrak{p}(x_r) \geqq 0$, also $k[x_1, \ldots, x_r] < R_\mathfrak{p}$. Ist $\nu_\mathfrak{p}(x_i^{-1}) > 0$, so kann, weil x_i^{-1} ein Primelement von $k[x_i^{-1}, x_i^{-1}x_1, \ldots, x_i^{-1}x_r]$ ist, nur $\nu_\mathfrak{p}(x_i^{-1}) = 1$ gelten. Dann muß aber $\nu_\mathfrak{p}(x_i^{-1}x_j) = 0$ sein $(j=1, \ldots, r)$, weil der Restklassenkörper von $R_\mathfrak{p}$ den Transzendenzgrad $r-1$ hat. Es folgt somit $\nu_\mathfrak{p}(x_i^{-1}) = 1$ für $j=1, \ldots, r$. $R_\mathfrak{p}$ ist daher der lokale Ring von (x_1^{-1}) in $k[x_1^{-1}, x_1^{-1}x_2, \ldots, x_1^{-1}x_r]$.

Außerdem haben wir gesehen, daß stets gleichzeitig $\nu_\mathfrak{p}(x_i) \geqq 0$ ist $(i=1, \ldots, r)$ oder $\nu_\mathfrak{p}(x_i) = -1$ $(i=1, \ldots, r)$, was $\mathfrak{n}_{x_1} = \cdots = \mathfrak{n}_{x_r}$ beweist.

Wir bezeichnen den gemeinsamen Nennerdivisor von $x_1, \ldots, x_r$ mit $\mathfrak{n}_x$. Er ist, wie aus dem Beweis von Hilfssatz 7 hervorgeht, ein Primdivisor 1. Art auf V. $k_0 = k$ ist ein Differentialkonstantenkörper von V. Wir bestimmen den Divisor von $dx_1 \wedge \cdots \wedge dx_r$ auf V. Es ergibt sich

$$(dx_1 \wedge \cdots \wedge dx_r) = \frac{1}{\mathfrak{n}_x^{r+1}}.$$

Für alle Primdivisoren $\mathfrak{p}$, die von $\mathfrak{n}_x$ verschieden sind, ist $x_1, \ldots, x_r$ ein System uniformisierender Parameter, daher $(dx_1 \wedge \cdots \wedge dx_r)_\mathfrak{p} = 1$. Für $\mathfrak{p} = \mathfrak{n}_x$ ist $x_1^{-1}, x_1^{-1}x_2, \ldots, x_1^{-1}x_r$ ein System uniformisierender Parameter. Wegen

$$dx_1 \wedge \cdots \wedge dx_r = -x_1^{r+1} \cdot dx_1^{-1} \wedge d(x_1^{-1}x_2) \wedge \cdots \wedge d(x_1^{-1}x_r)$$

folgt nun sofort die Behauptung. Zusammen mit Satz 10 ergibt sich

SATZ 11. *Der Erweiterungsdivisor von $\mathfrak{n}_x$ auf V' sei ebenfalls mit $\mathfrak{n}_x$ bezeichnet. Dann besteht die kanonische Klasse von V' aus allen Divisoren*

$$\frac{\mathfrak{d}\left(\frac{V'}{V}\right)}{\mathfrak{n}_x^{r+1}},$$

wo $\mathfrak{d}$ eine bezüglich einer nichttrivialen Linearform σ von L über K gebildete σ-Differente von V' über V ist.

§ 6. Anwendungen auf algebraische Funktionenkörper einer Variablen

Es sei L ein algebraischer Funktionenkörper einer Variablen über k und $K = k(x)$, wo x ein transzendentes Element von L über k ist. Wir wenden Satz 11 auf die singularitätenfreien vollständigen Modelle V von K und V' von L an, d.h. auf die aus allen Bewertungsringen von K über k bzw. von L über k bestehenden Modelle. Da die Voraussetzungen von Satz 11 erfüllt sind, folgt, daß die kanonische Klasse von V' aus allen Divisoren $\mathfrak{d}(V'/V) \cdot \mathfrak{n}_x^{-2}$

besteht, wo $\mathfrak{d}$ eine bezüglich einer nichttrivialen Linearform von L über K gebildete Differente ist. Da in $\mathfrak{d}(V'/V) = \prod_{\mathfrak{P}} \mathfrak{d}(S_{\mathfrak{P}}/R_{\mathfrak{p}})$ alle Bewertungsringe $S_{\mathfrak{P}}$ von L über k auftreten, stimmt diese Klasse bekanntlich mit der kanonischen Klasse von L überein, die im Riemann-Rochschen Satz auftritt.

Satz 12. *Ist K ein algebraischer Funktionenkörper einer Variablen über k und V das singularitätenfreie vollständige Modell von K/k, so stimmt die kanonische Klasse von V mit der kanonischen Klasse von K/k, die im Riemann-Rochschen Satz auftritt, überein. Das Geschlecht von K ist das geometrische Geschlecht von V.*

Es soll jetzt noch gezeigt werden, daß man bei Verwendung von Differentialformen von K/k statt von K über einem Differentialkonstantenkörper k_0 nicht zur kanonischen Klasse von K/k zu gelangen braucht, wenn man die in § 4 angegebene „natürliche" Divisordefinition verwendet. Man kann dies schon für separabel erzeugbare Funktionenkörper K einer Variablen über k zeigen, vorausgesetzt, daß k unvollkommen ist. Es sei x ein separierendes transzendentes Element von K/k. Dann ist $M(K/k) = K\,dx$. Nach der Formel von F. K. Schmidt aus § 4 ergibt sich für den Divisor von dx

$$(dx) = \frac{\prod_{\mathfrak{p}} \mathfrak{d}_0\left(\frac{R_{\mathfrak{p}}}{k[x^{e_{\mathfrak{p}}}]}\right)}{\mathfrak{n}_x^2 \cdot \prod_{\mathfrak{p}} \mathfrak{d}_1\left(\frac{R_{\mathfrak{p}}}{k}\right)}.$$

$\mathfrak{p}$ durchläuft alle Primdivisoren von K über k. Aus der Separabilität von K über $k(x)$ folgt, daß $\prod_{\mathfrak{p}} \mathfrak{d}_0\left(\frac{R_{\mathfrak{p}}}{k[x^{e_{\mathfrak{p}}}]}\right)$ gleich der Dedekindschen Differente $\mathfrak{d}_x$ von K über $k(x)$ ist, also $(dx) = \frac{\mathfrak{d}_x}{\mathfrak{n}_x^2} \cdot \frac{1}{\prod_{\mathfrak{p}} \mathfrak{d}_1\left(\frac{R_{\mathfrak{p}}}{k}\right)}$. Da der Faktor $\prod_{\mathfrak{p}} \mathfrak{d}_1\left(\frac{R_{\mathfrak{p}}}{k}\right)$ von 1 verschieden sein kann, sobald k unvollkommen ist (vgl. [4], Satz 3), folgt, daß (dx) kein Divisor der kanonischen Klasse von K zu sein braucht.

Literatur

[1] Cartan, H., u. C. Chevalley: Schémas normaux; Morphismes; Ensembles constructibles. Sem. Cartan-Chevalley **8** (1955/56).

[2] Chevalley, C.: Les schémas (I, II). Sem. Cartan-Chevalley **8** (1955/56).

[3] Chevalley, C.: Introduction to the Theory of Algebraic Functions of one Variable. New York 1951.

[4] Kunz, E.: Die Primidealteiler der Differenten in allgemeinen Ringen. J. reine angew. Math. **204**, 165–182 (1960).

[5] Lamprecht, E.: Über s-Differenten und Differentiale algebraischer Funktionenkörper einer Variablen. Arch. Math. **4**, 412–424 (1953).

[6] Tate, J.: Genus Change in inseparable Extensions of Function Fields. Proc. Amer. Math. Soc. **3**, 400–406 (1952).

[7] Zariski, O.: Introduction to the Theory of Algebraic Surfaces. Vorlesungsausarbeitung, Harvard Univ. 1957/58.

Heidelberg, Mathematisches Institut der Universität

(Eingegangen am 11. Oktober 1960)

NASTOLD, H.-J.
Math. Zeitschr. 76, 75–84 (1961)

Zum Dualitätssatz in inseparablen Funktionenkörpern der Dimension 1

Von

HANS-JOACHIM NASTOLD

Nachdem von E. KUNZ in der vorausgehenden Arbeit [4] gezeigt wurde, daß für eine algebraische Mannigfaltigkeit V mit inseparablem Funktionenkörper die Differentiation über einem geeigneten Unterkörper des Konstantenkörpers das richtige Tangentialbündel und die richtigen Differentialformen auf V liefert und daß insbesondere im Fall der Dimension 1 die so gebildeten Differentialformen der höchsten Stufe die richtige kanonische Klasse ergeben, stellte sich die Aufgabe, mit diesen Differentialen den *Dualitätssatz* zu beweisen. Dies soll im folgenden für den Fall der Dimension 1, also für Kurven auch mit inseparablem Funktionenkörper durchgeführt werden. Wir verlassen hier also den Standpunkt von A. Weils Foundations, wo der Funktionenkörper einer Varietät stets separabel ist. Über die Beziehung des *Dualitätssatzes* zum *Riemann-Rochschen Satz* und seine Verallgemeinerung auf höhere Dimensionen sei hier auf [7], insbesondere chap. II, verwiesen. Der Inhalt des Dualitätssatzes besteht in der Präzisierung einer *Dualität zwischen dem Raum der Repartitionen* einerseits (vgl. dazu [1]) *und dem Raum der durch einen Differentiationsprozeß gewonnenen* und daher rechnerisch leichter zu handhabenden *Differentialformen höchster Stufe* andererseits.

Das die Dualität liefernde Skalarprodukt definieren wir mit Hilfe von *Residuen* der Formen höchster Stufe. Der Begriff des Residuums wird dabei folgendermaßen gewonnen: Durch eine *Spurbildung* wird die Form in einen rationalen Unterkörper heruntergedrückt. Lediglich für Formen dieses rationalen Funktionenkörpers ist nun noch ein Residuum zu erklären. Dies geschieht mit Hilfe der Partialbruchzerlegung. Der Beweis des *Residuensatzes* beschränkt sich nun auf den Nachweis dieses Satzes im rationalen Funktionenkörper und wird somit trivial. Die Untersuchung der *Divisoren* der auftretenden Formen erfordert die Kenntnis des Zusammenhangs zwischen Spuren, Differenten und Differentiation.

Die eben genannte *Spurbildung* $L \to K$, L/K endlich algebraisch, für Formen höchster Stufe definieren wir sogleich für Funktionenkörper beliebiger Dimension. Sie dürfte auch an sich von Interesse sein. Wir erhalten sie durch Zusammensetzung und für einen p-Schritt, $L^p \subset K$, $[L:K]=p$, mit Hilfe der von J. TATE in [8] angegebenen Elementspur. Während die letztere aber von der Auswahl einer p-Basis abhängt, erweist sich unsere Spur für Differentialformen als invariant.

Grundlegend für unsere Entwicklungen ist der Begriff des *Differentialkonstantenkörpers*, so wie er in [4] definiert wurde, sowie dessen Eigenschaften.

Wir schließen uns daher auch in den Bezeichnungen eng an [4] an. So bedeutet $M(K/k_0)$, $M(R/k_0)$ usw. stets den Differentialmodul von K, R usw. über $k_0 \cdot M^t(K/k_0) = \wedge^t M(K/k_0)$ dessen t-te äußere Potenz. Mit Ausnahme des letzten Abschnitts 7. beschränken wir uns auf den Fall von Funktionenkörpern der Charakteristik $p > 0$.

Auf einen entsprechenden Dualitätssatz für Kurven mit Singularitäten und für Formen der komplementären Stufen r und $t-r$ kommen wir bei anderer Gelegenheit zurück.

Herrn E. KUNZ danke ich für wertvolle Diskussionen und die Überlassung des Manuskriptes seiner noch unveröffentlichten Arbeit [4].

1. Spur für Differentialformen höchster Stufe

$K = k(x_1, \ldots, x_n)$ sei ein algebraischer Funktionenkörper der Dimension r über k, V ein *Modell* von K/k, L eine endlich algebraische Erweiterung von K. Wir betrachten eine Kette von Zwischenkörpern $K \subset K_0 \subset K_1 \subset \cdots \subset K_s = L$, K_0 separable Hülle von K in L, $K_i^p \subset K_{i-1}$, $[K_i : K_{i-1}] = p$ für $i \geqq 1$, und in K_i jeweils das von V abgeleitete normale Modell V_i, $V_s = W$. $k_0 \subset k$ sei ein gemeinsamer wie in [4], § 5, c gewählter Differentialkonstantenkörper für die Modelle V, V_i, W. Mit $[K_i : K_i^p(k_0)] = [K : K^p(k_0)] = p^t$ definieren wir eine *Spurabbildung für Differentialformen der höchsten Stufe* t $S_{L/K} : M^t(L/k_0) \to M^t(K/k_0)$ durch Hintereinanderausführung von Abbildungen $S_{K_i/K_{i-1}} : M^t(K_i/k_0) \to M^t(K_{i-1}/k_0)$ und $S_{K_0/K} : M^t(K_0/k_0) \to M^t(K/k_0)$. Es genügt also, Spuren zu definieren für die beiden Fälle a) $L^p \subset K$, $[L : K] = p$, und b) L/K separabel algebraisch.

a) Wie in [4], § 5, c) ist $K^p(k_0) \subset L^p(k_0) \subset K \subset L$. Wir wählen eine p-Basis $x_1, \ldots, x_{t-1}$ von K über $L^p(k_0)$ und eine p-Basis x von L über K. Dann ist x^p eine p-Basis von $L^p(k_0)$ über $K^p(k_0)$ und $x_1, \ldots, x_{t-1}, x$ bzw. $x^p, x_1, \ldots, x_{t-1}$ sind p-Basen von L über $L^p(k_0)$ bzw. K über $K^p(k_0)$, ihre Differentiale also Basen der Differentialmoduln $M(L/k_0)$ bzw. $M(K/k_0)$. Wir definieren nun die Spur der beliebigen Form $\lambda \cdot dx_1 \wedge \cdots \wedge dx_{t-1} \wedge dx \in M^t(L/k_0)$, $\lambda \in L$, durch die Formel $S_{L/K}(\lambda \cdot dx_1 \wedge \cdots \wedge dx_{t-1} \wedge dx) = \sigma_x(\lambda) \cdot dx^p \wedge dx_1 \wedge \cdots \wedge dx_{t-1} \in M^t(K/k_0)$, wo $\sigma_x : L \to K$ die Tatesche Spur gebildet mit dem Element x bezeichnet [8]. Die so definierte Spur ist unabhängig von den gewählten p-Basen $x_1, \ldots, x_{t-1}$, und x. Ersetzt man zunächst x durch eine andere p-Basis y von L über K so wird $\lambda \cdot dx_1 \wedge \cdots \wedge dx_{t-1} \wedge dx = \lambda D_y x \cdot dx_1 \wedge \cdots \wedge dx_{t-1} \wedge dy$, wenn $dx = (D_y x) \cdot dy + a_1 dx_1 + \cdots + a_{t-1} dx_{t-1}$, $a_i \in L$, und man hat nun $S_{L/K}(\lambda D_y x \cdot dx_1 \wedge \cdots \wedge dx_{t-1} \wedge dy) = \sigma_y(\lambda D_y x) \cdot dy^p \wedge dx_1 \wedge \cdots \wedge dx_{t-1} = \sigma_x\left(\lambda D_y x \cdot (D_y x)^{p-1}\right) \cdot dy^p \wedge dx_1 \wedge \cdots \wedge dx_{t-1} = \sigma_x(\lambda) \cdot D_y x^p \cdot dy^p \wedge dx_1 \wedge \cdots \wedge dx_{t-1} = \sigma_x(\lambda) \cdot dx^p \wedge dx_1 \wedge \cdots \wedge dx_{t-1}$. Ersetzt man $x_1, \ldots, x_{t-1}$ durch eine andere p-Basis $y_1, \ldots, y_{t-1}$ von K über $L^p(k_0)$, so tritt zu λ ein Faktor aus K, der sich aus der Spur σ_x herausziehen läßt.

Ist $\mathfrak{q}$ ein Primdivisor 1. Art auf W, S der zugehörige Bewertungsring von L und $R = S \cap K$, so ist $S \cap L^p(k_0) = S^p[k_0]$ und $R \cap K^p(k_0) = R^p[k_0]$. Es gibt dann stets eine p-Basis y von L über K, so daß $S = R[y]$, und eine p-Basis $y_1, \ldots, y_{t-1}$ von K über $L^p(k_0)$, so daß $R = S^p[k_0, y_1, \ldots, y_{t-1}]$. y^p ist dann

eine p-Basis von $L^p(k_0)$ über $K^p(k_0)$ mit $S^p[k_0]=R^p[k_0, y^p]$ und $y_1, \ldots, y_{t-1}, y$ bzw. $y^p, y_1, \ldots, y_{t-1}$ sind p-Basen von L über $L^p(k_0)$ bzw. K über $K^p(k_0)$ mit $S=S^p[k_0, y_1, \ldots, y_{t-1}, y]$ bzw. $R=R^p[k_0, y^p, y_1, \ldots, y_{t-1}]$. Die Potenzprodukte der $y_1, \ldots, y_{t-1}, y$ bzw. $y^p, y_1, \ldots, y_{t-1}$ bilden Ganzheitsbasen der Bewertungsringe S bzw. R und ihre Differentiale Basen der Differentialmoduln $M(S/k_0)$ bzw. $M(R/k_0)$. Wir können dann mit diesen speziellen p-Basen die Spur der Form $\omega=\lambda' \cdot dy_1 \wedge \cdots \wedge dy_{t-1} \wedge dy$ so berechnen: $S_{L/K}(\omega)=\sigma_y(\lambda') \cdot dy^p \wedge dy_1 \wedge \cdots \wedge dy_{t-1}$. Nach Definition des Divisors einer Form ω, [4], § 3, ist $v_\mathfrak{q}(\omega)=v_\mathfrak{q}(\lambda')$, ebenso $v_\mathfrak{p}(S_{L/K}(\omega))=v_\mathfrak{p}(\sigma_y(\lambda'))$, wenn $\mathfrak{p}$ den zu R gehörenden Primdivisor 1. Art auf V bezeichnet. Da die Elemente $1, y, \ldots, y^{p-1}$ eine Basis von S über R bilden, ist der Differentenexponent $m_{\mathfrak{q}, \sigma_y}$ für die Linearform σ_y gleich 0, denn für die mit der Tateschen Spur σ_y gebildeten Differente $\mathfrak{D}$ gilt nach [4], § 5, b) $\mathfrak{D}=(D_y y)^{1-p} \cdot S=(1)$. Vgl. dazu auch [5], 2. und [1], S. 69ff.

b) Ist L/K separabel algebraisch, so wählen wir eine p-Basis $x_1, \ldots, x$ von K über $K^p(k_0)$. Diese ist dann zugleich p-Basis von L über $L^p(k_0)$, und wir definieren unsere Spurabbildung folgendermaßen $S_{L/K}(\lambda \cdot dx_1 \wedge \cdots \wedge dx_t)=\sigma_{L/K}(\lambda) \cdot d'x_1 \wedge \cdots \wedge d'x_t$, wo $\sigma_{L/K}$ die gewöhnliche Spur für separabel algebraische Erweiterungen, d die Differentiation $L \to M(L/k_0)$ und d' die Differentiation $K \to M(K/k_0)$ bedeuten. Wiederum hängt die so definierte Abbildung nicht von der gewählten p-Basis $x_1, \ldots, x_t$ ab. Für einen Primdivisor 1. Art $\mathfrak{q}$ auf W, den zugehörigen Bewertungsring S von L und $R=S \cap K$ kann man wieder eine p-Basis $y_1, \ldots, y_t$ von K über $K^p(k_0)$ mit $R=R^p[k_0, y_1, \ldots, y_t]$ finden, deren Differentiale eine Basis von $M(R/k_0)$ bilden, und $S_{L/K}$ in der Form $S_{L/K}(\lambda' \cdot dy_1 \wedge \cdots \wedge dy_t)=\sigma_{L/K}(\lambda') \cdot d'y_1 \wedge \cdots \wedge d'y_t$ berechnen.

Insgesamt ist die so definierte Spur eine K-lineare Abbildung

$$S_{L/K}: M^t(L/k_0) \to M^t(K/k_0).$$

2. Lokale Spuren

Wir definieren durch eine leichte Modifikation neben unserer globalen Spur für jeden Primdivisor 1. Art $\mathfrak{q}$ auf W eine *lokale Spur* $S_{L/K,\mathfrak{q}}$, indem wir dieselbe Kette von Zwischenkörpern $K \subset K_0 \subset K_1 \subset \cdots \subset K_s=L$ wie in 1. zugrundelegen, die Abbildung $S_{L/K,\mathfrak{q}}$ wieder durch Hintereinanderausführung der Abbildung $S_{K_i/K_{i-1},\mathfrak{q}}$ erklären, $S_{K_i/K_{i-1},\mathfrak{q}}$ unverändert gleich $S_{K_i/K_{i-1}}$ setzen und lediglich die Definition der Spur für den separablen Anteil K_0/K abändern: Wir setzen für L/K separabel algebraisch sonst wie unter 1. b) $S_{L/K,\mathfrak{q}}(\lambda \cdot dx_1 \wedge \cdots \wedge dx_t)=\sigma_{L/K,\mathfrak{q}}(\lambda) \cdot dx_1 \wedge \cdots \wedge dx_t$, wo $\sigma_{L/K,\mathfrak{q}}$ die gewöhnliche lokale Spur, d.h. die Spur $\sigma_{L_\mathfrak{q}/K_\mathfrak{p}}: L_\mathfrak{q} \to K_\mathfrak{p}$ der Komplettierung $L_\mathfrak{q}$ von L nach dem Primdivisor $\mathfrak{q}$ in die Komplettierung $K_\mathfrak{p}$ von K nach dem Primdivisor $\mathfrak{p}$, über dem $\mathfrak{q}$ liegt $(\mathfrak{q} \to \mathfrak{p})$, bedeutet. λ ist dabei vermöge der kanonischen Einbettung $L \to L_\mathfrak{q}$ als Element von $L_\mathfrak{q}$ aufzufassen, und $\sigma_{L/K,\mathfrak{q}}(\lambda)$ ist aus $K_\mathfrak{p}$. Man erhält so eine Abbildung $S_{L/K,\mathfrak{q}}: M^t(L/k_0) \to K_\mathfrak{p} \underset{K}{\otimes} M^t(K/k_0)$. Wegen $\sum_{\mathfrak{q} \to \mathfrak{p}} \sigma_{L/K,\mathfrak{q}}(\lambda)=\sigma_{L/K}(\lambda)$ gilt $\sum_{\mathfrak{q} \to \mathfrak{p}} S_{L/K,\mathfrak{q}}(\omega)=S_{L/K}(\omega) \in M^t(K/k_0)$. Diese Formel gilt für eine

beliebige endlich algebraische Erweiterung L/K: Es folgt aus der Formel für den separablen Anteil K_0/K $\sum_{\mathfrak{q}\to\mathfrak{p}} S_{K_0/K,\mathfrak{q}}(\omega') = S_{K_0/K}(\omega')$ mit $\omega' = S_{L/K_0}(\omega) =$ $S_{L/K_0,\mathfrak{q}}(\omega)$ $\sum_{\mathfrak{q}\to\mathfrak{p}} S_{K_0/K,\mathfrak{q}}\left(S_{L/K_0,\mathfrak{q}}(\omega)\right) = S_{K_0/K}\left(S_{L/K_0}(\omega)\right)$, also nach Definition der Spuren

$$\sum_{\mathfrak{q}\to\mathfrak{p}} S_{L/K,\mathfrak{q}}(\omega) = S_{L/K}(\omega). \tag{1}$$

Die lokalen Spuren sind K-lineare Abbildungen

$$S_{L/K,\mathfrak{q}}: M^t(L/k_0) \to K_{\mathfrak{p}} \underset{K}{\otimes} M^t(K/k_0),\ \mathfrak{q}\to\mathfrak{p}.$$

3. Residuen

Von nun an beschränken wir uns auf einen Funktionenkörper L der Dimension 1 über k. Wir wählen in L einen rationalen Unterkörper $K=k(x)$ und eine Erzeugende x von K/k aus. Alle auftretenden Funktionenkörper der Dimension 1 seien mit dem eindeutig bestimmten vollständigen normalen, also singularitätenfreien Modell versehen. Letzteres besteht gerade aus allen Bewertungsringen mit diesem Quotientenkörper, die k enthalten. k_0 sei wieder ein wie in 1. gewählter gemeinsamer Differentialkonstantenkörper: $k_0 \subset k$. L/K ist dann endlich algebraisch, und wir haben in 1. und 2. Spurabbildungen $L\to K$ für Differentialformen höchster Stufe definiert. Zur Definition des *Residuums* $\mathrm{res}_{\mathfrak{q}}(\omega)$ einer solchen Form $\omega\in M^t(L/k_0)$ an der Stelle $\mathfrak{q}$ von L genügt es daher, das Residuum $\mathrm{res}_{\mathfrak{p}}(\pi)$ einer Form $\pi\in M^t(K/k_0)$ an einer Stelle $\mathfrak{p}$ von K zu definieren. Dies geschieht im Anschluß an [2] mit Hilfe der Partialbruchzerlegung im rationalen Funktionenkörper K: Wir wählen noch eine p-Basis $x_1, \ldots, x_{t-1}$ von k über $k^p(k_0)$. Dann ist $x_1, \ldots, x_{t-1}, x$ eine p-Basis von K über $K^p(k_0)$. Jede Form $\pi\in M^t(K/k_0)$ läßt sich in der Gestalt $\pi=\alpha\cdot dx_1\wedge\cdots\wedge dx_{t-1}\wedge dx$, $\alpha\in K$, schreiben. Ihr ordnen wir die Form $\alpha\cdot dx\in M(K/k)$ zu und definieren für eine Stelle $\mathfrak{p}$ mit $v_{\mathfrak{p}}(x)\geqq 0$, wenn die rationale Funktion $\alpha\in K$ in x die Partialbruchentwicklung

$$\alpha = \frac{c_{-\varepsilon}x^{s-1}+\cdots+c_{-\varepsilon}^{(s-1)}}{p(x)^{\varepsilon}} + \frac{c_{-\varepsilon+1}x^{s-1}+\cdots+c_{-\varepsilon+1}^{(s-1)}}{p(x)^{\varepsilon-1}} + \cdots + \frac{c_{-1}x^{s-1}+\cdots+c_{-1}^{(s-1)}}{p(x)} + \beta,$$

$\beta\in K$, $v_{\mathfrak{p}}(\beta)\geqq 0$, $p(x)$ das zu $\mathfrak{p}$ gehörige normierte Primpolynom aus $k[x]$, Grad $p(x)=s$, besitzt,

$$\mathrm{res}_{\mathfrak{p}}(\pi) = \mathrm{res}_{\mathfrak{p}}(\alpha\cdot dx) = c_{-1}. \tag{2}$$

Ist $v_{\mathfrak{p}}(x)<0$, so ersetzen wir x durch $y=1/x$ und definieren das Residuum genau so mit der Partialbruchzerlegung in y. Wie man mühelos nachrechnet, hängt $\mathrm{res}_{\mathfrak{p}}(\pi)$ nicht von der Auswahl von x ab. Wir unterdrücken diese Rechnung, da wir die Unabhängigkeit von x im folgenden nicht benötigen, denken uns also ein erzeugendes Element x von K/k ein für allemal fest gewählt. Andererseits hängt $\mathrm{res}_{\mathfrak{p}}(\pi)$ ab von der Auswahl der p-Basis $x_1, \ldots, x_{t-1}$

von k über $k^p(k_0)$. Bei Ersetzung durch eine andere p-Basis $y_1, \ldots, y_{t-1}$ von k über $k^p(k_0)$ multiplizieren sich alle $\mathrm{res}_\mathfrak{p}(\pi)$ mit ein und demselben Faktor aus dem Konstantenkörper k. Das Residuum hat die Eigenschaften: (i) $\mathrm{res}_\mathfrak{p}: M^t(K/k_0) \to k$ ist eine k-lineare Abbildung (ii) $\mathrm{res}_\mathfrak{p}(\pi) = 0$ für $v_\mathfrak{p}(\alpha) \geqq 0$.

Nunmehr können wir das Residuum für ein lokales Differential $\pi_\mathfrak{p} \in K_\mathfrak{p} \underset{K}{\otimes} M^t(K/k_0)$ definieren: Mit $\pi_\mathfrak{p} = \alpha_\mathfrak{p} \cdot dx_1 \wedge \cdots \wedge dx_{t-1} \wedge dx$, $\alpha_\mathfrak{p} \in K_\mathfrak{p}$, approximieren wir $\alpha_\mathfrak{p}$ durch ein $\alpha \in K$, so daß $v_\mathfrak{p}(\alpha_\mathfrak{p} - \alpha) \geqq 0$, und setzen für $\pi = \alpha \cdot dx_1 \wedge \cdots \wedge dx_{t-1} \wedge dx$ $\mathrm{res}_\mathfrak{p}(\pi_\mathfrak{p}) = \mathrm{res}_\mathfrak{p}(\pi)$. Zufolge (i) und (ii) ist $\mathrm{res}_\mathfrak{p}(\pi_\mathfrak{p})$ von der Auswahl des approximierenden Elementes $\alpha \in K$ unabhängig.

Mit Hilfe der lokalen Spuren $S_{L/K,\mathfrak{q}}: M^t(L/k_0) \to K_\mathfrak{p} \underset{K}{\otimes} M^t(K/k_0)$, $\mathfrak{q} \to \mathfrak{p}$, und der Residuen lokaler Differentiale des rationalen Funktionenkörpers K $\mathrm{res}_\mathfrak{p}: K_\mathfrak{p} \underset{K}{\otimes} M^t(K/k_0) \to k$ definieren wir das Residuum $\mathrm{res}_\mathfrak{q}(\omega)$ einer Form $\omega \in M^t(L/k_0)$ an der Stelle $\mathfrak{q}$ von L durch die Formel

$$\mathrm{res}_\mathfrak{q}(\omega) = \mathrm{res}_\mathfrak{p}\big(S_{L/K,\mathfrak{q}}(\omega)\big), \qquad \mathfrak{q} \to \mathfrak{p}. \tag{3}$$

Die Abbildung $\mathrm{res}_\mathfrak{q}: M^t(L/k_0) \to k$ ist k-linear. Aus (1) und (3) ergibt sich sofort $\sum\limits_{\mathfrak{q} \to \mathfrak{p}} \mathrm{res}_\mathfrak{q}(\omega) = \sum\limits_{\mathfrak{q} \to \mathfrak{p}} \mathrm{res}_\mathfrak{p}\big(S_{L/K,\mathfrak{q}}(\omega)\big) = \mathrm{res}_\mathfrak{p}\big(\sum\limits_{\mathfrak{q} \to \mathfrak{p}} S_{L/K,\mathfrak{q}}(\omega)\big) = \mathrm{res}_\mathfrak{p}\big(S_{L/K}(\omega)\big)$, also

$$\sum_{\mathfrak{q} \to \mathfrak{p}} \mathrm{res}_\mathfrak{q}(\omega) = \mathrm{res}_\mathfrak{p}\big(S_{L/K}(\omega)\big). \tag{4}$$

Für das Folgende grundlegend sind

HILFSSATZ 1. *Ist* $v_\mathfrak{q}(\omega) \geqq 0$, *so ist* $\mathrm{res}_\mathfrak{q}(\omega) = 0$,

und

HILFSSATZ 2. *Es gibt ein* $\omega \in M^t(L/k_0)$ *mit* $v_\mathfrak{q}(\omega) = -1$ *und* $\mathrm{res}_\mathfrak{q}(\omega) \neq 0$.

Den Beweis dieser Hilfssätze stellen wir zurück bis zum Schluß. Wir beweisen zunächst den Residuensatz und definieren dann mit Hilfe der Residuen ein Skalarprodukt $\langle \omega, \mathfrak{r} \rangle$ zwischen Differentialen $\omega \in M^t(L/k_0)$ und Repartitionen $\mathfrak{r} \in R$, das uns den Dualitätssatz liefern wird. Nachträglich ergibt sich so, daß unsere Definition von $\mathrm{res}_\mathfrak{q}(\omega)$ in folgendem Sinne eindeutig ist: Andere Wahl der in die Definition von $\mathrm{res}_\mathfrak{q}(\omega)$ eingehenden willkürlichen Bestimmungsstücke, etwa des rationalen Zwischenkörpers K usw. liefert eventuell andere Residuen, die sich jedoch, falls der Konstantenkörper k in L algebraisch abgeschlossen ist, nur um ein und denselben festen Faktor $\varkappa$ aus k unterscheiden: $\mathrm{res}_\mathfrak{q}^{(1)}(\omega) = \varkappa \cdot \mathrm{res}_\mathfrak{q}^{(2)}(\omega)$, $\varkappa \in k$, für alle ω und $\mathfrak{q}$.

4. Der Residuensatz

SATZ 1. *Für jede Form* $\omega \in M^t(L/k_0)$ *ist*

$$\sum_{\mathfrak{q}} \mathrm{res}_\mathfrak{q}(\omega) = 0. \tag{5}$$

$\mathfrak{q}$ *durchläuft dabei alle Primdivisoren von* L *oder alle Punkte der zugehörigen Kurve.*

BEWEIS. Nach Hilfssatz 1 sind in der Summe (5) nur endlich viele Summanden von 0 verschieden. Ferner ist nach (3) und (4) $\sum_{\mathfrak{q}} \operatorname{res}_{\mathfrak{q}}(\omega) = \sum_{\mathfrak{p}} \sum_{\mathfrak{q}\to\mathfrak{p}} \operatorname{res}_{\mathfrak{q}}(\omega) = \sum_{\mathfrak{p}} \operatorname{res}_{\mathfrak{p}}(S_{L/K}(\omega))$. Wegen $S_{L/K}(\omega) \in M^1(K/k_0)$ genügt es also, den Residuensatz für Formen des rationalen Funktionenkörpers zu beweisen. Dazu unterscheiden wir für eine feste Erzeugung $K = k(x)$ die Primdivisoren $\mathfrak{p}$ von K mit $\nu_{\mathfrak{p}}(x) \geqq 0$ von dem zur Gradbewertung bezüglich x gehörigen Primdivisor ∞. Für die Stelle ∞ hat man zur Bildung des Residuums x durch $y = 1/x$ zu ersetzen und dann in der Partialbruchzerlegung in y des Koeffizienten α' das Residuum zum Primpolynom $y \in k[y]$ aufzusuchen. Auf Grund der Linearität der Residuenbildung und der Partialbruchzerlegung in x der Funktionen $\alpha \in K$ genügt es, um $\sum_{\mathfrak{p}} \operatorname{res}_{\mathfrak{p}}(\omega) = 0$ für ein beliebiges $\omega \in M^1(K/k_0)$ nachzuweisen, diese Beziehung für die ω der Gestalt $\omega_1 = x^n \cdot dx_1 \wedge \cdots \wedge dx_{t-1} \wedge dx$, $n \geqq 0$, und $\omega_2 = \frac{c_{-\alpha} x^{s-1} + \cdots + c_{-\alpha}^{(s-1)}}{p(x)^{\alpha}} \cdot dx_1 \wedge \cdots \wedge dx_{t-1} \wedge dx$, $\alpha \geqq 1$, zu beweisen. Nun ist $\nu_{\mathfrak{p}}(x^n) \geqq 0$, also $\operatorname{res}_{\mathfrak{p}}(\omega_1) = 0$ für alle $\mathfrak{p}$. Ferner ist $\omega_1 = \frac{-1}{y^{n+2}} \cdot dx_1 \wedge \cdots \wedge dx_{t-1} \wedge dy$, also auch $\operatorname{res}_{\infty}(\omega_1) = 0$. Ebenso ist für $\alpha \geqq 2$ $\operatorname{res}_{\mathfrak{p}}(\omega_2) = 0$ für alle $\mathfrak{p}$ und $\omega_2 = \frac{z(x)}{p(x)^{\alpha}} \cdot dx_1 \wedge \cdots \wedge dx_{t-1} \wedge dx = \frac{-y^{\alpha s - r - 2} z'(y)}{p'(y)^{\alpha}} \cdot dx_1 \wedge \cdots \wedge dx_{t-1} \wedge dy$, wo $r = \operatorname{Grad} z(x) \leqq s-1$, $s = \operatorname{Grad} p(x)$, also $\alpha s - r - 2 \geqq (\alpha - 1)s - 1 \geqq 0$. $z'(y)$ und $p'(y)$ sind dabei die Polynome, die aus $z(x)$ und $p(x)$ entstehen, wenn man $x = 1/y$ setzt und die Nenner beseitigt. Da $p'(y)$ ein Primpolynom in $k[y]$ ist, ist $\nu_{\infty}(\omega_2) \geqq 0$, also $\operatorname{res}_{\infty}(\omega_2) = 0$. Schließlich ist für $\alpha = 1$ für den zu $p(x)$ gehörigen Primdivisor $\mathfrak{p}_0$ $\operatorname{res}_{\mathfrak{p}_0}(\omega_2) = c_{-1}$ und $\operatorname{res}_{\mathfrak{p}}(\omega_2) = 0$ für $\mathfrak{p} \neq \mathfrak{p}_0$. Hier ist $\alpha s - r - 2 \geqq s - (s-1) - 2 = -1$ und nur im Falle $r = s - 1 < 0$. Für diesen Fall wird $\omega_2 = \frac{-z'(y)}{p'(y)\, y} \cdot dx_1 \wedge \cdots \wedge dx_{t-1} \wedge dy = \left(\frac{z_1(y)}{p'(y)} + \frac{-c_{-1}}{y} \right) \cdot dx_1 \wedge \cdots \wedge dx_{t-1} \wedge dy$, also $\operatorname{res}_{\infty}(\omega_2) = -c_{-1}$.

5. Der Dualitätssatz

Im Riemann-Rochschen Satz in der auf A. WEIL zurückgehenden Form (vgl. dazu [1], [6] und insbesondere [7], chap. II) sind die *Repartitionen* oder Bewertungsvektoren von Bedeutung. Eine Repartition $\mathfrak{r}$ ist eine Familie $\{r_{\mathfrak{q}}\}_{\mathfrak{q} \in W}$ von Elementen $r_{\mathfrak{q}} \in L$ mit $\nu_{\mathfrak{q}}(r_{\mathfrak{q}}) \geqq 0$ für fast alle $\mathfrak{q}$, wo W die zum Funktionenkörper L gehörige Kurve, $\mathfrak{q}$ die Primdivisoren oder Punkte von W bezeichnen. Die Repartitionen bilden eine Algebra über k. Für einen Divisor D von L bezeichne $R(D)$ den Unterraum von R bestehend aus den $\mathfrak{r} = \{r_{\mathfrak{q}}\}$ mit $\nu_{\mathfrak{q}}(r_{\mathfrak{q}}) \geqq -\nu_{\mathfrak{q}}(D)$. Die Zuordnung $f \in L \to \{r_{\mathfrak{q}}\}$ mit $r_{\mathfrak{q}} = f$ für alle $\mathfrak{q} \in W$ liefert eine Injektion von L in R. Wir bezeichnen mit $I(D)$ den k-Vektorraum $R/(R(D) + L)$, mit $J(D)$ den Raum der k-linearen Formen auf $I(D)$: $J(D) = I(D)'$. Für $D' \geqq D$ ist $R(D') > R(D)$, also $J(D) > J(D')$. Der Raum $J = \bigcup_D J(D)$ ist ein L-Vektorraum der Dimension 1: Für $f \in L$ und $\alpha \in J$ soll $f\alpha$ die k-lineare Abbildung $\mathfrak{r} \to \langle \alpha, f\mathfrak{r} \rangle$ sein.

Um an die übliche Bezeichnung anzuschließen, schreiben wir hier für $M^1(L/k_0)$, den *Raum der Differentialformen höchster Stufe*, Ω^1. Ω^1 ist ein L-

Vektorraum der Dimension 1. Wir werden im folgenden einen L-Isomorphismus zwischen Ω^1 und J stiften. $(\omega)=\sum_{\mathfrak{q}\in W}\nu_{\mathfrak{q}}(\omega)\cdot\mathfrak{q}$ sei der Divisor von $\omega\in\Omega^1$. Für einen Divisor D bezeichnen wir mit $\Omega^1(D)$ den k-Vektorraum bestehend aus 0 und den Formen $\omega\neq 0$ mit $(\omega)\geqq D$.

Nun definieren wir ein *Skalarprodukt* $\langle\omega,\mathfrak{r}\rangle$ zwischen Differentialformen $\omega\in\Omega^1$ und Repartitionen $\mathfrak{r}\in R$:

$$\langle\omega,\mathfrak{r}\rangle=\sum_{\mathfrak{q}\in W}\operatorname{res}_{\mathfrak{q}}(\mathfrak{r}_{\mathfrak{q}}\omega). \tag{6}$$

Da $\nu_{\mathfrak{q}}(\mathfrak{r}_{\mathfrak{q}}\omega)\geqq 0$ für fast alle $\mathfrak{q}$, hat die Summe (6) nach Hilfssatz 1 nur endlich viele von 0 verschiedene Summanden. $\langle\omega,\mathfrak{r}\rangle$ ist wegen der k-Linearität von $\operatorname{res}_{\mathfrak{q}}$ ein Skalarprodukt zwischen den k-Vektorräumen Ω^1 und R mit den Eigenschaften:

(i) $\langle\omega,\mathfrak{r}\rangle=0$ für $\mathfrak{r}\in L$ nach dem Residuensatz.

(ii) $\langle\omega,\mathfrak{r}\rangle=0$ für $\mathfrak{r}\in R(D)$ und $\omega\in\Omega^1(D)$, denn dann ist $\nu_{\mathfrak{q}}(\mathfrak{r}_{\mathfrak{q}}\omega)\geqq 0$ für alle $\mathfrak{q}\in W$, und die Behauptung folgt aus Hilfssatz 1.

(iii) Für $f\in L$ ist $\langle f\omega,\mathfrak{r}\rangle=\langle\omega,f\mathfrak{r}\rangle$.

Bezeichnen wir für $\omega\in\Omega^1$ mit $\Theta(\omega)$ die durch $\mathfrak{r}\to\langle\omega,\mathfrak{r}\rangle$ definierte Linearform auf R, so ergibt sich aus (i) und (ii): Die Zuordnung $\omega\to\Theta(\omega)$ bildet $\Omega^1(D)$ in $J(D)$, den dualen Raum von $R/(R(D)+L)$, ab. Sie ist nach (iii) L-linear.

Satz 2 (Dualitätssatz). *Für jeden Divisor D ist Θ ein k-Isomorphismus von $\Omega^1(D)$ auf $J(D)$. $\langle\omega,\mathfrak{r}\rangle$ stiftet also eine Dualität zwischen den Vektorräumen $\Omega^1(D)$ und $I(D)=R/(R(D)+L)$.*

Beweis. Wir zeigen zunächst, daß aus $\Theta(\omega)\in J(D)$ $\omega\in\Omega^1(D)$ folgt. Wäre das nicht der Fall, so gäbe es ein $\mathfrak{q}\in W$ mit $\nu_{\mathfrak{q}}(\omega)<\nu_{\mathfrak{q}}(D)$, also $\nu_{\mathfrak{q}}(\omega)+1\leqq\nu_{\mathfrak{q}}(D)$. Nach Hilfssatz 2 existiert ein $\omega'\in\Omega^1$ mit $\nu_{\mathfrak{q}}(\omega')=-1$ und $\operatorname{res}_{\mathfrak{q}}(\omega')\neq 0$. Es ist $\omega'=\lambda\cdot\omega$ mit $\lambda\in L$, und $\nu_{\mathfrak{q}}(\lambda)=-\nu_{\mathfrak{q}}(\omega)-1\geqq-\nu_{\mathfrak{q}}(D)$. Für die Repartition $\mathfrak{r}$ mit $\mathfrak{r}_{\mathfrak{p}}=0$ für $\mathfrak{p}\neq\mathfrak{q}$ und $\mathfrak{r}_{\mathfrak{q}}=\lambda$ gilt $\mathfrak{r}\in R(D)$, und es ist $\operatorname{res}_{\mathfrak{q}}(\mathfrak{r}_{\mathfrak{q}}\omega)=\operatorname{res}_{\mathfrak{q}}(\omega')\neq 0$, dagegen $\operatorname{res}_{\mathfrak{p}}(\mathfrak{r}_{\mathfrak{p}}\omega)=\operatorname{res}_{\mathfrak{p}}(0)=0$ für $\mathfrak{p}\neq\mathfrak{q}$, also $\langle\omega,\mathfrak{r}\rangle=\sum_{\mathfrak{q}\in W}\operatorname{res}_{\mathfrak{q}}(\mathfrak{r}_{\mathfrak{q}}\omega)\neq 0$. $\Theta(\omega)$ sollte aber nach Voraussetzung auf $R(D)$ verschwinden. — Daraus folgt: Θ ist monomorph. Denn aus $\Theta(\omega)=0$ ergibt sich $\omega\in\Omega^1(D')$ für jeden Divisor D', also $\omega=0$. Da $\Theta:\Omega^1\to J$ ein L-Monomorphismus und Ω^1, J beides L-Vektorräume der Dimension 1 sind, ist Θ ein L-Isomorphismus von Ω^1 auf J. Zu $\alpha\in J(D)$ gibt es daher ein $\omega\in\Omega^1$ mit $\Theta(\omega)=\alpha$ und $\omega\in\Omega^1(D)$.

Korollar 1. *Ist k in L algebraisch abgeschlossen und sind Θ_1, Θ_2 zwei wie oben gewonnene L-Isomorphismen $\Omega^1\to J$, so ist $\Theta_1(\omega)=\varkappa\cdot\Theta_2(\omega)$ mit festem $\varkappa\in k$ für alle $\omega\in\Omega^1$. Für die zugehörigen Residuen gilt ebenfalls* $\operatorname{res}_{\mathfrak{q}}^{(1)}(\omega)=\varkappa\cdot\operatorname{res}_{\mathfrak{q}}^{(2)}(\omega)$.

BEWEIS. Durch Θ_1, Θ_2 wird ein L-Isomorphismus φ: $J \to J$ induziert. Es ist also $\Theta_1(\omega) = f \cdot \Theta_2(\omega)$ mit festem $f \in L$. Aus Satz 2 ergibt sich aber $\nu_\mathfrak{q}(f) \geqq 0$ für alle $\mathfrak{q}$, also $f = \varkappa \in k$. Mit $\alpha = \Theta(\omega)$ gilt aber, wenn r die Repartition mit $r_\mathfrak{q} = 1$, $r_\mathfrak{p} = 0$ für $\mathfrak{p} \neq \mathfrak{q}$ bezeichnet, für das zugehörige Residuum nach (6) $\operatorname{res}_\mathfrak{q}(\omega) = \langle \alpha, r \rangle$, also $\operatorname{res}_\mathfrak{q}^{(1)}(\omega) = \varkappa \cdot \operatorname{res}_\mathfrak{q}^{(2)}(\omega)$ für alle $\omega \in \Omega^1$.

Einer Linearform $\alpha \in J$ läßt sich ein Divisor zuordnen: $(\alpha) = \operatorname{Max}\{D | \alpha \in J(D)\}$.

KOROLLAR 2. *Es ist* $(\omega) = (\Theta(\omega))$.

Da nach dem Riemann-Rochschen Satz die Klasse der Divisoren (α), $\alpha \in J$, eindeutig bestimmt ist, gilt dasselbe von der Klasse der Divisoren (ω), $\omega \in \Omega^1 = M^1(L/k_0)$. Während also Ω^1 noch von der Wahl des Differentialkonstantenkörpers k_0 abhängt, hat sich somit ergeben:

KOROLLAR 3. *Die Klasse der Divisoren* (ω), $\omega \in \Omega^1 = M^1(L/k_0)$, *ist von der Wahl des Differentialkonstantenkörpers* k_0 *unabhängig.*

Dies läßt sich auch direkt und für Funktionenkörper beliebiger Dimension beweisen [4], § 3.

BEMERKUNG. Unser Dualitätssatz, der eine Dualität zwischen Formen der Stufen 0 und t (t höchste Stufe) herstellt, ist Spezialfall eines allgemeineren Dualitätssatzes, der eine entsprechende Beziehung zwischen Formen der komplementären Stufen r und $t - r$ herstellt. Wir werden an anderer Stelle darauf zurückkommen.

6. Nachweis der Hilfssätze 1 und 2

HILFSSATZ 1. Wir beweisen zunächst mit den Bezeichnungen von 1. für endlich algebraisches L/K und $\omega \in M^1(L/k_0)$: Aus $\nu_\mathfrak{q}(\omega) \geqq 0$ folgt $\nu_\mathfrak{p}(S_{L/K,\mathfrak{q}}(\omega)) \geqq 0$, $\mathfrak{q} \to \mathfrak{p}$. Nach Definition der Spur genügt es, dies für einen p-Schritt L/K (1. a)) und einen separabel algebraischen Schritt L/K (1. b)) jeweils einzeln zu zeigen. a) Mit den in 1. a) zu $\mathfrak{q}$ und S gewählten speziellen Basen $y_1, \ldots, y_{t-1}, y$ bzw. $y^p, y_1, \ldots, y_{t-1}$ ist für $\omega = \lambda' \cdot dy_1 \wedge \cdots \wedge dy_{t-1} \wedge dy$ wegen $\nu_\mathfrak{q}(\omega) = \nu_\mathfrak{q}(\lambda') \geqq 0$ in $S_{L/K,\mathfrak{q}}(\omega) = \sigma_y(\lambda') \cdot dy^p \wedge dy_1 \wedge \ldots \wedge dy_{t-1}$ $\nu_\mathfrak{p}(\sigma_y(\lambda')) \geqq 0$, $\mathfrak{q} \to \mathfrak{p}$, (nach dem in 1. a) über den Differentenexponenten von σ_y Bemerkten) und folglich $\nu_\mathfrak{p}(S_{L/K,\mathfrak{q}}(\omega)) \geqq 0$. — b) Wir wählen zunächst zu S, dem zum Primdivisor $\mathfrak{q}$ von L gehörigen Bewertungsring, eine p-Basis $x_1, \ldots, x_t$ von L über $L^p(k_0)$ mit $S = S^p[k_0, x_1, \ldots, x_t]$, deren Differentiale eine Basis von $M(S/k_0)$ bilden, und wie in 1. b) eine ebensolche p-Basis $y_1, \ldots, y_t$ zu $R = S \cap K$. Nun ist in

$$\omega = \lambda \cdot dx_1 \wedge \cdots \wedge dx_t \qquad \nu_\mathfrak{q}(\lambda) = \nu_\mathfrak{q}(\omega) \geqq 0$$

und

$$\omega = \lambda \cdot \frac{\partial(x_1 \ldots x_t)}{\partial(y_1 \ldots y_t)} \cdot dy_1 \wedge \cdots \wedge dy_t = \lambda \cdot \frac{\partial(y_1 \ldots y_t)^{-1}}{\partial(x_1 \ldots x_t)} \cdot dy_1 \wedge \cdots \wedge dy_t.$$

Da $M(S/R) = M(S/k_0)/\langle S\,dR \rangle$, ist $\frac{\partial(y_1 \ldots y_t)}{\partial(x_1 \ldots x_t)} \cdot S = \mathfrak{D}_0(S/R)$, der 0-ten Kählerschen Differente von S über R. Diese ist aber gleich der Dedekindschen Differente von S über R. Aus dem bekannten Zusammenhang zwischen lokaler Spur $S_{L/K,\mathfrak{q}}$ und Differentenexponent (vgl. dazu [5], 2. u. [1], S. 69ff.)

ergibt sich mit $S_{L/K,\mathfrak{q}}(\omega)=\sigma_{L/K,\mathfrak{q}}\left(\lambda\,\frac{\partial(y_1\dots y_t)^{-1}}{\partial(x_1\dots x_t)}\right)\cdot d'y_1\wedge\dots\wedge d'y_t$ aus

$$v_{\mathfrak{q}}\left(\lambda\,\frac{\partial(y_1\dots y_t)^{-1}}{\partial(x_1\dots x_t)}\right)\geqq -\,v_{\mathfrak{q}}\left(\mathfrak{D}_0\left(\frac{S}{R}\right)\right)$$

$$v_{\mathfrak{p}}\left(\sigma_{L/K,\mathfrak{q}}\left(\lambda\,\frac{\partial(y_1\dots y_t)^{-1}}{\partial(x_1\dots x_t)}\right)\right)\geqq 0,\quad\text{also } v_{\mathfrak{p}}\left(S_{L/K,\mathfrak{q}}(\omega)\right)\geqq 0,\qquad \mathfrak{q}\to\mathfrak{p}.$$

Daraus ergibt sich nun Hilfssatz 1 mit den Bezeichnungen und Definitionen von 3. folgendermaßen: Für $\omega\in M^1(L/k_0)$ mit $v_{\mathfrak{q}}(\omega)\geqq 0$ ist wegen $v_{\mathfrak{p}}\left(S_{L/K,\mathfrak{q}}(\omega)\right)\geqq 0$, $\mathfrak{q}\to\mathfrak{p}$, in der Darstellung von $S_{L/K,\mathfrak{q}}(\omega)$ durch die wie eben speziell zu $\mathfrak{p}$ in K gewählte p-Basis $y_1,\dots,y_t$ $S_{L/K,\mathfrak{q}}(\omega)=\varkappa\cdot dy_1\wedge\dots\wedge dy_t$, $\varkappa\in K_{\mathfrak{p}}$ (den Strich bei d' lassen wir nun wieder weg) $v_{\mathfrak{p}}(\varkappa)\geqq 0$. Mit der in 3. benutzten p-Basis $x_1,\dots,x_{t-1}$, x von K über $K^p(k_0)$ wird dann für $v_{\mathfrak{p}}(x)\geqq 0$ $S_{L/K,\mathfrak{q}}(\omega)=\varkappa\cdot dy_1\wedge\dots\wedge dy_t=\varkappa\,\frac{\partial(x_1\dots x_{t-1}x)^{-1}}{\partial(y_1\dots y_t)}\cdot dx_1\wedge\dots\wedge dx_{t-1}\wedge dx$. Wie eben in b) ist aber

$$\frac{\partial(x_1\dots x_{t-1}x)}{\partial(y_1\dots y_t)}\cdot R=\mathfrak{D}_0\left(\frac{R}{k_0[x_1\dots x_{t-1}][x]}\right)=\mathfrak{D}_0\left(\frac{R}{k[x]}\right),$$

der 0-ten Kählerschen Differente von R über $k[x]$. Letztere ist aber gleich (1), da der Bewertungsring R von K ein Quotientenring von $k[x]$ ist. Somit ist auch in $S_{L/K,\mathfrak{q}}(\omega)=\alpha\cdot dx_1\wedge\dots\wedge dx_{t-1}\wedge dx$ $v_{\mathfrak{p}}(\alpha)\geqq 0$ und folglich nach 3. $\operatorname{res}_{\mathfrak{q}}(\omega)=0$.

Hilfssatz 2. Wir zeigen zunächst: Es gibt ein $\omega\in M^1(L/k_0)$ mit $v_{\mathfrak{q}}(\omega)=-1$ und $v_{\mathfrak{p}}\left(S_{L/K,\mathfrak{q}}(\omega)\right)=-1$, $\mathfrak{q}\to\mathfrak{p}$. Wie im Beweis von Hilfssatz 1 ergibt sich dies für einen p-Schritt und einen separablen Schritt jeweils aus der Bedeutung des Differentenexponenten. Durch Zusammensetzung erhält man daraus die Behauptung: Für $K_{i+1}\supset K_i\supset K_{i-1}$ und Primdivisoren $\mathfrak{q}_{i+1}\to\mathfrak{q}_i\to\mathfrak{q}_{i-1}$ ist $S_{K_{i+1}/K_{i-1},\mathfrak{q}_{i+1}}=S_{K_i/K_{i-1},\mathfrak{q}_i}\circ S_{K_{i+1}/K_i,\mathfrak{q}_{i+1}}$. Es gibt nun ein $\omega\in M^1(K_{i+1}/k_0)$ mit $v_{\mathfrak{q}_{i+1}}(\omega)=-1$ und $v_{\mathfrak{q}_i}\left(S_{K_{i+1}/K_i,\mathfrak{q}_{i+1}}(\omega)\right)=-1$ und ein $\pi\in M^1(K_i/k_0)$ mit $v_{\mathfrak{q}_i}(\pi)=-1$ und $v_{\mathfrak{q}_{i-1}}\left(S_{K_i/K_{i-1},\mathfrak{q}_i}(\pi)\right)=-1$. Nun sind $S_{K_{i+1}/K_i,\mathfrak{q}_{i+1}}(\omega)$, $\pi\in M^1(K_i/k_0)$ und haben beide den $\mathfrak{q}_i$-Wert -1. Folglich ist $\pi=\varkappa\cdot S_{K_{i+1}/K_i,\mathfrak{q}_{i+1}}(\omega)$ mit $\varkappa\in K_i$ und $v_{\mathfrak{q}_i}(\varkappa)=0$. Dann ist $S_{K_{i+1}/K_i,\mathfrak{q}_{i+1}}(\varkappa\omega)=\pi$, also $v_{\mathfrak{q}_{i+1}}(\varkappa\omega)=-1$ und $v_{\mathfrak{q}_{i-1}}\left(S_{K_{i+1}/K_{i-1},\mathfrak{q}_{i+1}}(\varkappa\omega)\right)=-1$. Damit ergibt sich nun Hilfssatz 2: Es gibt zunächst ein $\omega\in M^1(L/k_0)$ mit $v_{\mathfrak{q}}(\omega)=-1$ und $v_{\mathfrak{p}}\left(S_{L/K,\mathfrak{q}}(\omega)\right)=-1$, $\mathfrak{q}\to\mathfrak{p}$. Wie im Beweis von Hilfssatz 1 ist dann auch in $S_{L/K,\mathfrak{q}}(\omega)=\alpha\cdot dx_1\wedge\dots\wedge dx_t\wedge dx$ $v_{\mathfrak{p}}(\alpha)=-1$. Die Partialbruchzerlegung in x von α hat also, wenn $p(x)\in k[x]$ das zu $\mathfrak{p}$ gehörige normierte Primpolynom bezeichnet, die Gestalt $\alpha=\frac{z(x)}{p(x)}+\beta$, $v_{\mathfrak{p}}(\beta)\geqq 0$, $z(x)\in k[x]$, Grad $z(x)=r\leqq s-1$, $s=$ Grad $p(x)$. Wegen $x^{(s-1)-r}\alpha=\frac{(s-1)-r\,z(x)}{p(x)}+x^{(s-1)-r}\beta$, $v_{\mathfrak{p}}(x^{(s-1)-r}\beta)\geqq 0$, und Grad $x^{(s-1)-r}z(x)=s-1$ ist für $\omega=x^{(s-1)-r}\cdot\omega\in M^1(L/k_0)$ $S_{L/K,\mathfrak{q}}(\omega')=x^{(s-1)-r}\alpha\cdot dx_1\wedge\dots\wedge dx_{t-1}\wedge dx$, also $v_{\mathfrak{q}}(\omega')=-1$ und

$$\operatorname{res}_{\mathfrak{q}}(\omega')=\operatorname{res}_{\mathfrak{p}}\left(S_{L/K,\mathfrak{q}}(\omega')\right)\neq 0.$$

Bemerkung. Die beim Beweis der Hilfssätze 1 und 2 sich ergebenden Aussagen über $S_{L/K,\mathfrak{q}}$ sind offenbar unter den allgemeineren Voraussetzungen von 1., also für beliebige Dimensionen, richtig.

7. Der Fall separabler Funktionenkörper

Der Funktionenkörper L/k sei nun *separabel*, wobei auch der bisher ausgeschlossene Fall der *Charakteristik* 0 zugelassen werde. Wir betrachten dann den Differentialmodul $M(L/k)$ über dem arithmetischen Konstantenkörper k und anstelle von $M'(L/k_0)$ das über L eindimensionale $M(L/k)$ selbst. K sei ein rationaler Zwischenkörper, so daß L/K separabel algebraisch ist. Mit der in 1. b) und 2. definierten Spur und dem in 3. definierten Residuum beweist man nun unverändert den Residuensatz und definiert wie in 5. ein Skalarprodukt $\langle \omega, r \rangle$ zwischen $\omega \in \Omega = M(L/k)$ und $r \in R$. Die Zuordnung $\omega \to \Theta(\omega)$ liefert dann wie in 5. einen L-Isomorphismus zwischen Ω und J. Nicht übertragen läßt sich jedoch bei unvollkommenem k der auf die Divisoren bezügliche Teil unserer obigen Überlegungen. Dem entspricht die Feststellung in [4], § 6, daß die „natürliche" Divisordefinition aus [4], § 4 für Formen aus $M(L/k)$, also bei Differentiation über dem arithmetischem Konstantenkörper, nicht zur richtigen kanonischen Klasse führt. Daß sich trotzdem eine vernünftige Dualität auch mit den Differentialen aus $M(L/k)$ herstellen läßt, läßt den den Schluß zu, daß die in [4], § 4 angegebene Divisordefinition nicht das Richtige trifft. Und zwar ist dies genau für diejenigen Stellen $\mathfrak{q}$ mit zugehörigen Bewertungsringen $S_\mathfrak{q}$ der Fall, für die $M(S_\mathfrak{q}/k)$ *Torsion* hat, also *nicht* isomorph in $M(L/k)$ eingebettet ist. Nach [3], § 3 sind dies genau die Stellen $\mathfrak{q}$ mit $\mathfrak{D}_1(S_\mathfrak{q}/k) \neq (1)$, und das Produkt $\prod_\mathfrak{q} \mathfrak{D}_1(S_\mathfrak{q}/k)$ gibt in [4], § 6 auch gerade die Abweichung von der richtigen kanonischen Klasse an. Natürlich läßt sich im Falle der Charakteristik 0 das in 5. und 6. über Divisoren Bewiesene ebenfalls zeigen, und zwar im Wesentlichen unverändert wie in den Fällen b), wenn man nur immer anstelle der dortigen zu Primdivisoren $\mathfrak{q}$, $\mathfrak{p}$ speziell gewählten p-Basen Systeme uniformierender Parameter zu $\mathfrak{q}$, $\mathfrak{p}$ verwendet.

Literatur

[1] CHEVALLEY, C.: Introduction to the Theory of Algebraic Functions of one Variable. New York 1951.

[2] EICHLER, M.: Differentiale und Riemann-Rochscher Satz in algebraischen Funktionenkörpern einer Variablen. Hamb. Abh. **24**, 5—11 (1960).

[3] KUNZ, E.: Die Primidealteiler der Differenten in allgemeinen Ringen. J. reine angew. Math. **204**, 165—182 (1960).

[4] KUNZ, E.: Differentialformen inseparabler algebraischer Funktionenkörper. Math. Z. 76, 55—74 (1961).

[5] LAMPRECHT, E.: Über s-Differenten und Differentiale algebraischer Funktionenkörper einer Variablen. Arch. Math. **4**, 412—424 (1953).

[6] ROQUETTE, P.: Über den Riemann-Rochschen Satz in Funktionenkörpern vom Transzendenzgrad 1. Math. Nachr. **19**, 375—404 (1958).

[7] SERRE, J. P.: Groupes algébriques et corps de classes, Paris 1959.

[8] TATE, J.: Genus Change in Inseparabel Extensions of Function Fields. Proc. Amer. Math. Soc. **3**, 400—406 (1952).

Heidelberg, Mathematisches Institut der Universität

(Eingegangen am 12. Oktober 1960)

André, J.
Math. Zeitschr. 76, 85–102 (1961)

Über Parallelstrukturen*)

Teil I: Grundbegriffe

Von

Johannes André

Einleitung

1. Die Translationsebenen sind unter den affinen Ebenen dadurch gekennzeichnet, daß ihre Translationsgruppen transitiv auf den Punkten sind, d.h. daß es zu je zwei Punkten immer eine (und damit dann genau eine) Translation gibt, die den einen in den anderen Punkt abbildet. Durch die grundlegenden Untersuchungen von M. Hall (1943) u. a.[1]) hat man einen, wenn auch noch unvollständigen, Einblick in ihre Struktur gewonnen. Zur tieferen Einsicht scheint es zweckmäßig, die Translationsebenen dahingehend zu verallgemeinern, daß die Forderung der Transitivität der Translationsgruppe etwas abgeschwächt wird. Folgende Fälle in dieser Hinsicht wurden bisher untersucht: (a) Transitivität nur auf den Punkten einer Geraden, das ergibt Ebenen über den sog. cartesischen Gruppen [Baer (1942)]; (b) Existenz von nichtidentischen Translationen in jeder Richtung, und als weitere Verallgemeinerung davon (c) Existenz nichtidentischer Translationen längs mindestens zweier Richtungen. Diese letzten beiden Typen wurden erst kürzlich [André (1958), (1959)] eingeführt und T- bzw. t-Ebenen genannt. Zunächst erhebt sich die Frage, ob es überhaupt Ebenen mit diesen Eigenschaften gibt, die keine Translationsebenen sind. Das ist in der Tat der Fall: In André (1959) wurden Beispiele von T-Ebenen angegeben, die keine Translationsebenen sind[2]). Man gewinnt sie aus der (desarguesschen) Ebene über dem rationalen Zahlkörper durch geeignetes „Knicken" von Geraden, ähnlich wie es etwa Moulton (1902), Naumann (1954, a) und Lombardo-Radice (1954) durchgeführt haben. Auf ähnliche Weise kann man auch t-Ebenen finden, die keine T-Ebenen sind. Über die allgemeine Struktur solcher Gebilde ist damit natürlich nichts ausgesagt. Es ist z.B. nicht bekannt, ob es *endliche* T-Ebenen gibt, die keine Translationsebenen sind[3]). Bisher weiß man folgendes [André (1958), Zappa (1959)]: Die Gruppe aller zentralen Kollineationen (Translationen und Streckungen) muß in einer endlichen T-Ebene[4]), die keine Translationsebene ist,

*) Teil einer von der Naturwissenschaftlich-Philosophischen Fakultät der Technischen Hochschule Braunschweig angenommenen Habilitationsschrift.

[1]) Eine zusammenfassende Darstellung über die Translationsebenen (mit Literaturangaben) findet man in Pickert (1955), S. 199–220.

[2]) Weitere Beispiele von T-Ebenen, sogar mit topologischer Struktur, die keine Translationsebenen sind, findet man in Salzmann (1958).

[3]) Dagegen gibt es endliche t-Ebenen, die keine Translationsebenen sind; vgl. dazu Hughes (1957), Zappa (1959).

[4]) Dies gilt auch unter der Annahme, daß es sich nur um eine t-Ebene handelt.

die Menge der Punkte in mindestens *drei* Transitivitätsgebiete zerlegen. Außerdem muß eine solche Ebene mehr als 81 (eigentliche) Punkte besitzen. Weitergehende Aussagen hat man nicht.

2. Zur genaueren Untersuchung der Eigenschaften von t- oder T-Ebenen erweist es sich als zweckmäßig, allgemeinere geometrische Strukturen zu betrachten, die sog. Parallelstrukturen. Ihre systematische Untersuchung ist der Inhalt dieser und drei weiterer Arbeiten. Bei den Parallelstrukturen ist, wie der Name schon andeuten soll, die Parallelität von Geraden der zentrale Begriff: Eine affine Parallelstruktur (a. P). ist eine Menge von Punkten und Geraden derart, daß durch zwei verschiedene Punkte genau eine Gerade geht (somit zwei verschiedene Geraden höchstens einen Punkt gemeinsam haben) und daß gewisse Paare von Geraden als zueinander „parallel" gelten, wobei das Euklidische Parallelenaxiom gültig sei; ferner müssen genügend viele Punkte vorhanden sein (Reichhaltigkeitsaxiom), genauer soll es drei nicht auf einer Geraden gelegene Punkte geben und jede Gerade mindestens einen Punkt enthalten. Während bei den affinen Ebenen durch die geometrische Struktur die Parallelität der Geraden festgelegt ist (genau die sich nicht schneidenden Geraden sind dort zueinander parallel), muß bei den a. P. *definiert* werden, welche Geraden als parallel gelten sollen: Zwei verschiedene parallele Geraden schneiden sich zwar nicht, es kann aber, im Gegensatz zu Ebenen, auch nichtparallele Geraden ohne gemeinsamen Punkt geben (sog. windschiefe Geraden).

Man kann leicht Beispiele von a. P. bilden, die keine Ebenen sind. Das einfachste dieser Art ist ein n-dimensionaler affiner Raum mit $n>2$, dessen eindimensionale Unterräume die Geraden unserer a. P. sind und deren Parallelität im Raum sich auf die a. P. überträgt.

Ganz entsprechend, wie man von einer affinen zur projektiven Ebene gelangt, kommt man durch Einführung uneigentlicher Elemente von einer a. P. zu einer projektiven Parallelstruktur (p. P.), die man auch axiomatisch kennzeichnen kann (s. § 2). Im Verlauf dieser und der weiteren Arbeiten ist aber hauptsächlich von a. P. die Rede.

3. Die vorliegende Note bringt die nötigen Grundbegriffe, während die weiteren drei Teile der Untersuchung spezieller Parallelstrukturen gewidmet sind, die in Analogie zu den Translationsebenen, t- bzw. T-Ebenen durch die Existenz genügend vieler Translationen gekennzeichnet sind. Zunächst bringt § 1 in Anlehnung an PICKERT (1955), Kap. 1.1, Begriff und Eigenschaften der Inzidenzstrukturen, soweit sie im folgenden benötigt werden. In § 2 werden dann die projektiven und affinen Parallelstrukturen definiert, und die oben skizzierte Beziehung zwischen beiden hergeleitet. Durch Dualisierung (d. h. Umwandlung von Punkten in Geraden und umgekehrt bei Erhaltung der Inzidenz) speziellerer Parallelstrukturen gelangt man zu einem andersgearteten Typ geometrischer Strukturen, den wohlbekannten Geweben [vgl. z. B. BLASCHKE-BOL (1938), PICKERT (1955), REIDEMEISTER (1930)]. In § 3 wird dieser Zusammenhang behandelt.

In § 4 werden die Kollineationen eingeführt und einige ihrer speziellen Typen wie Translationen und Streckungen genauer untersucht; alle diese

Begriffe werden ganz analog wie bei Ebenen definiert. Die Menge aller Kollineationen bildet (bei üblicher Einführung der Multiplikation) wieder eine Gruppe. Dagegen ist keineswegs sicher (und für den allgemeinen Fall weder bewiesen noch widerlegt), ob auch die Translationen eine Gruppe bilden. Die Schwierigkeit liegt darin, zu zeigen, daß auch das Produkt zweier Translationen keinen Punkt festläßt, oder aber die Identität ist. Der Beweis für die affinen Ebenen [BAER (1942); PICKERT (1955), S. 65] läßt sich auf den allgemeinen Fall nicht übertragen.

Noch einen Unterschied gibt es gegenüber den bekannten Sätzen in Ebenen: Während dort eine Translation mit einer Geraden auch jede zu ihr parallele (als Ganzes) festläßt, d.h. ein (uneigentliches) „Zentrum" besitzt (oder „zentral" ist), braucht dies bei Parallelstrukturen nicht der Fall zu sein, Beispiele hierfür werden in Teil II angegeben. Wir werden uns aber hauptsächlich (in den Teilen III und IV sogar ausschließlich) mit Parallelstrukturen beschäftigen, in denen jede Translation zentral ist; in solchen Strukturen bilden übrigens die Translationen eine Gruppe, was genau so wie bei Ebenen bewiesen werden kann. Dieser Satz kann auch noch für endliche Parallelstrukturen bewiesen werden; dies gelingt aber bisher nur mit Hilfe der Theorie der Gruppencharaktere. In einem abschließenden § 5 wird dieser Zusammenhang gebracht.

4. Für diese und die folgenden Noten über Parallelstrukturen werden durchweg folgende einheitliche Bezeichnungen eingeführt: Mit kleinen lateinischen Buchstaben werden — neben den natürlichen Zahlen — Punkte gekennzeichnet, mit großen lateinischen Buchstaben Mengen von Punkten, insbesondere Geraden (die hier im allgemeinen als Menge der auf ihr liegenden Punkte gedeutet werden), mit großen Frakturbuchstaben schließlich Mengen von Punktmengen, speziell Geradenmengen. Abbildungen, insbesondere Kollineationen, und Mengen von solchen werden mit kleinen bzw. großen griechischen Buchstaben bezeichnet. Große Schreibschriftlettern stehen für Inzidenzstrukturen, während die kleinen für Inzidenzstrukturen mit Parallelitätsrelationen, in diesen Noten für affine Parallelstrukturen, vorbehalten sind.

§ 1. Inzidenzstrukturen

1. Vorgegeben seien zwei nicht leere Mengen P und $\mathfrak{G}$, deren Elemente wir *Punkte* bzw. *Geraden* nennen wollen. Zwischen ihnen sei eine Relation I, genannt *Inzidenz* festgelegt, d.h. es soll eindeutig feststehen, ob ein Punkt $p \in P$ mit einer Geraden $G \in \mathfrak{G}$ in Inzidenz steht oder nicht. Im ersten Fall sagen wir „p inzidiert mit G", „p ist mit G inzident", oder auch „p liegt in G", „G enthält p", oder „G geht durch p"; wir schreiben dafür $p\,\mathrm{I}\,G$. Die Negation der Inzidenz („p inzidiert nicht mit G") bezeichnen wir mit $p\,\overline{\mathrm{I}}\,G$.

DEFINITION. Ein Paar $\mathfrak{I} = (P, \mathfrak{G}, \mathrm{I})$ nicht leerer Mengen P und $\mathfrak{G}$ mit einer Inzidenzrelation I zwischen deren Elementen nennen wir *Inzidenzstruktur*[5]), wenn gilt:

[5]) Von M. HALL in (1943), S. 233 als "partial plane" eingeführt. Den Namen "Inzidenzstruktur" hat G. PICKERT [(1955), S. 2ff.] zuerst verwendet.

(i) *Aus den vier Inzidenzen* $p_i I G_k$ $(p_i \in P,\ G_k \in \mathfrak{G};\ i, k = 1, 2)$ *folgt* $p_1 = p_2$ *oder* $G_1 = G_2$.

Das besagt, daß durch zwei verschiedene Punkte höchstens eine Gerade geht, und zwei verschiedene Geraden höchstens einen Punkt gemeinsam haben.

Beispiele von Inzidenzstrukturen sind beliebige nicht leere Untermengen von Punkten und Geraden einer projektiven Ebene [PICKERT (1955), S. 7] mit den in ihr gegebenen Inzidenzen. Dies ergibt übrigens bereits alle Inzidenzstrukturen: M. HALL [(1943), S. 236–244] bettete nämlich jede Inzidenzstruktur vermöge einer Art „freien Erweiterung" in eine projektive Ebene ein.

2. Statt $(P, \mathfrak{G}, I)$ schreiben wir auch kürzer $(P, \mathfrak{G})$, wenn kein Mißverständnis zu befürchten ist. Die Mengen P und $\mathfrak{G}$ sind in $\mathcal{I} = (P, \mathfrak{G})$ völlig gleichberechtigt. Statt $p I G$ $(p \in P,\ G \in \mathfrak{G})$ schreiben wir auch $p \in G$. Man kann nämlich die Inzidenz im allgemeinen als mengentheoretische Enthaltenseinsrelation $\in$ deuten: Inzidiert jede Gerade G einer Inzidenzstruktur $\mathcal{I}$ mit mindestens zwei Punkten [$\mathcal{I}$ heißt in diesem Falle *regulär*, vgl. PICKERT (1955), S. 2], so können wir G auffassen als Menge der mit ihr inzidierenden Punkte; nach (i) ist G durch diese Menge eindeutig bestimmt. Diesen Standpunkt werden wir im folgenden häufig mit Vorteil einnehmen[6]). Ausdrücklich sei jedoch vermerkt, daß wir das Zeichen $\in$ (anstatt I) für die Inzidenz eines Punktes mit einer Geraden auch bei nicht regulären Inzidenzstrukturen verwenden werden.

3. Liegen zwei verschiedene Punkte p und q einer Inzidenzstruktur $\mathcal{I}$ auf einer Geraden (und damit wegen Axiom (i) auf genau einer), so sagen wir auch, p sei mit q *verbindbar*; die gemeinsame Gerade nennen wir die *Verbindungs*gerade von p mit q und bezeichnen sie mit $\langle p, q \rangle_{\mathcal{I}}$ oder auch einfach mit $\langle p, q \rangle$, wenn kein Mißverständnis zu befürchten ist. Es ist $\langle p, q \rangle = \langle q, p \rangle$. Existiert keine durch p und q gehende Gerade, ist also p mit q nicht verbindbar, so ist $\langle p, q \rangle$ nicht definiert[7]).

Haben zwei verschiedene Geraden G und G' einen (und dann genau einen) Punkt gemeinsam, so sagen wir auch, G und G' *schneiden* sich; den gemeinsamen Punkt $G \cap G'$ bezeichnen wir als *Schnittpunkt* von G mit G'. Schneiden sich G und G' nicht und ist $G \neq G'$, so ist $G \cap G'$ leer: $G \cap G' = \emptyset$.

4. Die beiden folgenden Begriffe sind für das Weitere wichtig.

DEFINITION. Eine Gerade einer Inzidenzstruktur heißt *Hauptgerade*, wenn sie mit *jeder* anderen Geraden dieser Struktur einen (und damit genau einen) Punkt gemeinsam hat.

---

[6]) Da Punkte und Geraden gleichberechtigt sind [wegen ihrer symmetrischen Stellung in Axiom (i)], könnten wir natürlich dieselbe Auffassung hinsichtlich der Geraden einnehmen (d.h. einen Punkt als Menge der durch ihn gehenden Geraden auffassen: Geradenbüschel). Das wird aber in dieser Arbeit nicht geschehen.

[7]) Wenn wir $\langle p, q \rangle$ schreiben, so soll damit immer stillschweigend vorausgesetzt sein, daß p mit q verbindbar sei. Insbesondere ist dann $p \neq q$.

Dual[8]) dazu wird der Begriff des Hauptpunktes eingeführt. Ein Punkt einer Inzidenzstruktur heißt somit *Hauptpunkt*, wenn er mit *jedem* anderen Punkt der Struktur verbindbar ist.

In projektiven Ebenen ist also jede Gerade Hauptgerade und jeder Punkt Hauptpunkt, und sie sind dadurch, abgesehen von einem Reichhaltigkeitsaxiom [das besagt, daß es vier Punkte gibt, von denen keine drei auf einer Geraden liegen, vgl. PICKERT (1955), S. 7, (18)], unter den Inzidenzstrukturen gekennzeichnet. In affinen Ebenen [PICKERT (1955), S. 10] ist zwar jeder Punkt Hauptpunkt, aber keine Gerade Hauptgerade.

§ 2. Projektive und affine Parallelstrukturen

1. DEFINITION. Eine *projektive Parallelstruktur* (p. P.) $\mathfrak{P} = (P, \mathfrak{G})$ ist eine Inzidenzstruktur mit den folgenden zusätzlichen Eigenschaften:

(p 1) *Jeder Punkt von $\mathfrak{P}$ ist Hauptpunkt.*

(p 2) *$\mathfrak{P}$ enthält mindestens eine Hauptgerade H.*

(p 3) *$\mathfrak{P}$ enthält außerhalb H mindestens drei Punkte, die nicht auf einer Geraden liegen* (d. h. ein nichtausgeartetes Dreieck).

(p 4) *Jede Gerade von $\mathfrak{P}$ enthält mindestens zwei Punkte.*

BEISPIEL EINER p. P. Die Punkte der gewünschten p. P. sind die Punkte eines n-dimensionalen projektiven Raumes, als Geraden nehmen wir einerseits eine Hyperebene H und andererseits die nicht in H gelegenen Geraden (des projektiven Raumes). Offensichtlich ist H eine Hauptgerade der Struktur (und zwar die einzige, wenn $n > 2$ ist). Auch die Gültigkeit der übrigen Axiome ist sofort einzusehen.

BEMERKUNGEN. Axiom (p 1) bedeutet: *Durch je zwei verschiedene Punkte von $\mathfrak{P}$ geht genau eine Gerade.* Wegen (p 2) kann man, wie gleich gezeigt wird, in einer aus $\mathfrak{P}$ abgeleiteten Inzidenzstruktur eine vernünftige Parallelitätsrelation einführen. Die Axiome (p 3) und (p 4) sind Reichhaltigkeitsaussagen. Aus (p 3) folgt, daß H mindestens drei Punkte enthält. Sind nämlich p_i $(i = 1, 2, 3)$ drei Punkte mit der in (p 3) geforderten Eigenschaft, so sind $H \cap \langle p_i, p_j \rangle$ $(i, j = 1, 2, 3;\ i \neq j)$ drei voneinander verschiedene auf H gelegene Punkte. Wir können jetzt (p 3) für *jede* Gerade von $\mathfrak{P}$ beweisen, also:

$\mathfrak{P}$ enthält außerhalb jeder Geraden mindestens drei nicht auf einer Geraden gelegenen Punkte (die aber eventuell teilweise auf H liegen können).

Sei nämlich G von H eine verschiedene Gerade. Voraussetzungsgemäß enthält H mindestens drei Punkte, also wenigstens zwei, p und q, außerhalb G. Nach (p 3) enthält $\mathfrak{P}$ ferner einen weder zu G noch zu H gehörigen Punkt r. Offenbar sind p, q und r Punkte mit der gewünschten Eigenschaft. Nun kann man ganz entsprechend wie eben schließen:

SATZ 2.1. *Jede Hauptgerade einer p. P. enthält mindestens drei Punkte.*

Beim Beweis wird natürlich die Hauptgeradeneigenschaft wesentlich benutzt, und somit muß (p 4), die Existenz von genügend vielen Punkten auf

[8]) Zum Begriff der Dualität und des Dualitätsprinzips s. etwa PICKERT (1955), S. 2/3.

jeder Geraden, extra gefordert werden. Daß (p 4) von den Axiomen (p 1) bis (p 3) tatsächlich unabhängig ist, erkennt man an dem folgenden einfachen Beispiel: Sei $\mathfrak{P}$ eine p.P. (im Sinne der oben angegebenen Definition), H eine Hauptgerade von $\mathfrak{P}$ und p ein Punkt von H. Wir erweitern jetzt $\mathfrak{P}$ auf folgende Weise zu einer Inzidenzstruktur $\overline{\mathfrak{P}}$: Die Punkte von $\mathfrak{P}$ und $\overline{\mathfrak{P}}$ sollen dieselben sein. Zu den Geraden von $\mathfrak{P}$ soll noch genau eine weitere hinzugenommen werden, die p, aber keinen anderen Punkt enthält. Offensichtlich erfüllt $\overline{\mathfrak{P}}$ die Axiome (p 1) bis (p 3) [wiederum mit H als Hauptgerade], jedoch nicht (p 4).

2. Für uns im folgenden wichtiger ist der Begriff der *affinen Parallelstruktur* (a.P.)[9], den wir jetzt aus demjenigen der p.P. entwickeln wollen: Sei H eine Hauptgerade der p.P. $\mathfrak{P}=(P, \mathfrak{G})$. Wir bilden durch Wegnahme der Geraden H und der auf ihr gelegenen Punkte eine neue Struktur $\mathfrak{p}=\mathfrak{P}_H=(P-H, \mathfrak{G}-\{H\})$ [die wir affine Parallelstruktur nennen wollen], ganz entsprechend, wie man aus einer projektiven Ebene durch Fortnahme einer Geraden eine affine Ebene erhält [s. z.B. PICKERT (1955), S. 9—11]: Die *Punkte* von $\mathfrak{p}$ sind die nicht in H liegenden Punkte von $\mathfrak{P}$. Die *Geraden* von $\mathfrak{p}$ sind alle von H verschiedenen Geraden von $\mathfrak{P}$. Ein Punkt $p \in P-H$ soll genau dann mit einer Geraden $G \in \mathfrak{G}-\{H\}$ in $\mathfrak{p}$ *inzidieren*, wenn p mit G in $\mathfrak{P}$ inzidiert[10]. Außerdem führen wir im Geradenbereich von $\mathfrak{p}$ eine Relation, genannt *Parallelität*, auf folgende Weise ein: Zwei Geraden G und G' von $\mathfrak{p}$ heißen genau dann zueinander *parallel*, wenn $G \cap H = G' \cap H$ in $\mathfrak{P}$ gilt[11]. Wir sagen in diesem Falle auch: „G ist zu G' parallel", „G ist Parallele zu G', oder „G und G' liegen parallel zueinander" und schreiben dafür $G \parallel G'$. Die so erhaltene Parallelitätsrelation erfüllt offensichtlich die drei Äquivalenzbedingungen; wir können also die Menge $\mathfrak{G}-\{H\}$ der Geraden von $\mathfrak{p}$ in Klassen paralleler Geraden einteilen.

Die so konstruierte Struktur $\mathfrak{p}$ hat die folgenden Eigenschaften:

(a 0) *$\mathfrak{p}$ ist eine Inzidenzstruktur.*

(a 1) *Jeder Punkt von $\mathfrak{p}$ ist Hauptpunkt.*

(a 2) *Es gibt eine Einteilung der Geraden von $\mathfrak{p}$ in Klassen paralleler derart, daß es zu jedem Punkt p und jeder Geraden G genau eine zu G parallele Gerade durch p gibt.* (Euklidisches Parallelenaxiom.)

(a 3) *$\mathfrak{p}$ enthält mindestens drei Punkte, die nicht auf einer Geraden liegen.*

(a 4) *Jede Gerade von $\mathfrak{p}$ enthält mindestens einen Punkt.*

BEISPIEL. Aus dem im Anschluß an die Definition der p.P. gebrachten Beispiel erhält man durch Entfernung von H folgende a.P.: Punkte, Geraden

[9] Eine ähnliche Struktur wurde von G. ZAPPA (1959) unter dem Namen „piano parziale con parallelismo" eingeführt.

[10] Hierbei ist zu beachten, daß $\mathfrak{p}$ nicht mehr regulär zu sein braucht, denn nach Herausnahme von H kann es Geraden mit nur einem Punkt geben. Auch in diesem Fall sollen aber (definitionsgemäß) verschiedene Geraden von $\mathfrak{P}$ nach Herausnahme von H zu verschiedenen Geraden von $\mathfrak{p}$ werden.

[11] Wegen der Hauptgeradeneigenschaft von H ist jedenfalls $G \cap H \neq \emptyset$ in $\mathfrak{P}$.

und Parallelität in der Struktur entsprechen den Punkten, Geraden und der Parallelität (im üblichen Sinne) eines affinen n-dimensionalen Raumes.

Bemerkung. Während die Struktur einer p.P. durch die Inzidenzen von Punkten und Geraden eindeutig festgelegt ist, genügt dies für die Bestimmung einer a.P. nicht; hierzu muß zusätzlich, wie in (a 2) dargelegt, die Parallelenrelation eingeführt werden.

Beweis von (a 0) bis (a 4). Als Unterstruktur [Pickert (1955), S. 3] von $\mathfrak{P}$ ist $\mathfrak{p}$ wieder eine Inzidenzstruktur, also gilt (a 0). Die Verbindungsgerade von p mit q ($p, q \in P-H$; $p \neq q$) existiert und stimmt mit derjenigen in $\mathfrak{P}$ überein [12]), mithin ist auch (a 1) erfüllt. Zum Beweis von (a 2) sei $p \in P-H$ und $G \in \mathfrak{G}-\{H\}$. Dann ist $\langle G \cap H, p\rangle$, aufgefaßt als Gerade von $\mathfrak{p}$, eine Parallele zu G durch p. Weitere Geraden mit dieser Eigenschaft gibt es nicht, weil $\mathfrak{P}$ eine Inzidenzstruktur ist. Die Reichhaltigkeitsaussagen (a 3) und (a 4) folgen sofort aus den entsprechenden Axiomen (p 3) und (p 4).

Die [nach (a 2) eindeutig bestimmte] zu G parallele Gerade durch p bezeichnen wir mit $\langle G \| p\rangle_{\mathfrak{p}}$ oder auch kurz mit $\langle G \| p\rangle$, wenn kein Mißverständnis zu befürchten ist. Aus (a 2) folgt sofort, daß zwei voneinander verschiedene parallele Geraden in $\mathfrak{p}$ keinen Punkt gemeinsam haben. Dagegen gilt die Umkehrung nicht: Es gibt a.P. mit Geraden, die sich nicht schneiden, ohne zueinander parallel zu sein (sie treten im letztgenannten Beispiel auf, wenn $n>2$ ist); Geraden mit dieser Eigenschaft nennen wir zueinander *windschief*. Die affinen Ebenen sind unter den a.P. dadurch ausgezeichnet, daß es in ihnen keine zueinander windschiefen Geraden gibt.

3. Wir wollen jetzt umgekehrt zeigen, daß man jede Struktur mit den Eigenschaften (a 0) bis (a 4) auf die oben beschriebene Weise (d.h. durch Fortnahme einer Hauptgeraden aus einer p.P.) erhalten kann. Wir sind dann berechtigt, *eine a.P. axiomatisch als eine Struktur mit den Eigenschaften* (a 0) *bis* (a 4) *zu erklären* [13]). Die jetzt durchzuführende Erweiterung einer a.P. zu einer p.P. geht wieder ebenso vor sich wie die einer affinen zu einer projektiven Ebene.

Sei also $\mathfrak{p}=(P, \mathfrak{G})$ eine Struktur mit den Eigenschaften (a 0) bis (a 4); wir wollen sie in eine p.P. $\mathfrak{p}^*=\mathfrak{P}$ einbetten in dem eben angedeuteten Sinne. Zu diesem Zwecke führen wir in $\mathfrak{p}$ zunächst „uneigentliche Punkte" ein, in ganz analoger Weise, wie dies bei affinen Ebenen geschieht [Pickert (1955), S. 11]. Als *uneigentlichen Punkt* $\hat{G}$ der Geraden $G \in \mathfrak{G}$ definieren wir die Menge aller in $\mathfrak{p}$ zu G parallelen Geraden. Somit ist $\hat{G}=\hat{G}'$ gleichbedeutend mit $G \| G'$.

Punkte von $\mathfrak{P}$ sind diejenigen von $\mathfrak{p}$ *(eigentliche Punkte)* und die eben definierten $\hat{G}$ mit $G \in \mathfrak{G}$ *(uneigentliche Punkte)*. *Geraden* von $\mathfrak{P}$ sind die von $\mathfrak{p}$ *(eigentliche Geraden)* und die Menge aller uneigentlichen Punkte von $\mathfrak{p}$ die *uneigentliche Gerade*). Mit einer eigentlichen Geraden G sollen $\hat{G}$ und

[12]) Die Gerade $\langle p, q\rangle_{\mathfrak{p}}$ inzidiert also mit genau allen nicht in H gelegenen Punkten von $\langle p, q\rangle_{\mathfrak{P}}$.

[13]) Entsprechend wie bei den p.P. kann man zeigen, daß (a 4) von den übrigen Axiomen unabhängig ist.

genau diejenigen eigentlichen Punkte inzidieren, die bereits in $\mathfrak{p}$ auf G liegen. In der uneigentlichen Geraden sollen genau alle uneigentlichen Punkte liegen.

Wir zeigen jetzt, daß die so erhaltene Struktur $\mathscr{P}$ eine p.P. ist: Die Verbindungsgerade zweier eigentlicher Punkte ist dieselbe wie in $\mathfrak{p}$. Der eigentliche Punkt p wird mit $\hat{G}$ nach (a 2) eindeutig durch $\langle G \| p \rangle$ verbunden. Nur die uneigentliche Gerade enthält zwei verschiedene uneigentliche Punkte. Damit ist (p 1) bewiesen (und auch, daß $\mathscr{P}$ eine Inzidenzstruktur ist). Die uneigentliche Gerade H ist eine Hauptgerade, denn jede Gerade $G \neq H$ hat (definitionsgemäß) mit H den Schnittpunkt $\hat{G}$. Also gilt auch (p 2). Axiom (p 3) folgt sofort aus (a 3). Jede eigentliche Gerade G enthält mindestens zwei Punkte, nämlich $\hat{G}$ und nach (a 4) noch mindestens einen eigentlichen. Die uneigentliche Gerade enthält sogar wenigstens drei Punkte; denn nach (a 3) gibt es mindestens drei eigentliche Punkte p_i $(i=1, 2, 3)$, die nicht auf einer Geraden liegen, und die uneigentlichen Punkte der drei Geraden $\langle p_i, p_j \rangle$ $(i, j=1, 2, 3)$ sind voneinander verschieden. Damit haben wir auch (p 4) gewonnen.

4. In Ziffer 2 sind wir, ausgehend von einer p.P. zu einer a.P. gelangt, in Ziffer 3 umgekehrt von einer a.P. zu einer p.P. Wir wollen jetzt zeigen, daß man durch Hintereinanderausführung dieser beiden Operationen wieder zur Ausgangsstruktur zurückkommt, d.h. wir wollen beweisen:

$$(\mathscr{P}_H)^* = \mathscr{P}, \tag{2.1}$$

$$(\mathfrak{p}^*)_H = \mathfrak{p}. \tag{2.2}$$

[Die Gl. (2.1) ist hierbei, wie gleich ausgeführt, im Sinne einer Identifizierung der Strukturen gemäß einem Isomorphismus zu verstehen.]

BEWEIS VON (2.1). Die nicht auf H gelegenen Punkte von $\mathscr{P}$ lassen wir bei dem zu konstruierenden Isomorphismus zwischen $\mathscr{P}$ und $(\mathscr{P}_H)^*$ sich selbst entsprechen. Ebenso sollen die von H verschiedenen Geraden identifiziert werden. Dem Punkt $q \in H$ ordnen wir in $(\mathscr{P}_H)^*$ den Punkt $\hat{Q}$ zu; hierbei ist Q irgendeine von H verschiedene Gerade von $\mathscr{P}$ [und damit auch von $(\mathscr{P}_H)^*$] durch q. Der Geraden H wird die Menge aller so erhaltenen Punkte $\hat{Q}$ zugeordnet. Man sieht sofort, daß die so beschriebene Abbildung einen Isomorphismus zwischen den beiden Strukturen vermittelt.

BEWEIS VON (2.2). Man erhält $\mathfrak{p}^*$ aus $\mathfrak{p}$ durch Hinzunahme von H und der darauf liegenden Punkte. Nimmt man diese aus $\mathfrak{p}^*$ wieder heraus, so kommt man zu $\mathfrak{p}$ zurück.

Die so gewonnene enge Beziehung zwischen den projektiven und affinen Parallelstrukturen verallgemeinert den bekannten Zusammenhang zwischen den projektiven und affinen Ebenen [vgl. PICKERT (1955), S. 7—12]. Auf diese Weise kann die Theorie der p.P. weitgehend auf die der a.P. zurückgeführt werden und umgekehrt. Wir werden uns im folgenden hauptsächlich mit a.P. beschäftigen.

5. Für spätere Untersuchungen benötigen wir den Begriff der *Unterstruktur* einer affinen Parallelstruktur. G. PICKERT hat in (1955), S. 3 eine Unterstruktur $\mathfrak{I}'$ der Inzidenzstruktur $\mathfrak{I}$ dadurch gekennzeichnet, daß die Punkt- und Geradenmenge von $\mathfrak{I}'$ Untermengen der entsprechenden Mengen von $\mathfrak{I}$ sein und die Inzidenzrelationen beim Übergang zur Unterstruktur erhalten bleiben sollen. Für die Unterstruktur $\mathfrak{p}'$ einer a.P. $\mathfrak{p}$ müssen wir außerdem noch fordern, daß auch die Parallelität erhalten bleibt. Mithin ist definitionsgemäß $\mathfrak{p}'=(P', \mathfrak{G}')$ eine Unterstruktur von $\mathfrak{p}=(P, \mathfrak{G})$, wenn gilt:

(a) $P' \subseteq P$,

(b) $\mathfrak{G}' \subseteq \mathfrak{G}$,

(c) *ist* $p \in P'$ *und* $G \in \mathfrak{G}'$, *so inzidiert* p *mit* G *in* $\mathfrak{p}'$ *genau dann, wenn dies in* $\mathfrak{p}$ *der Fall ist,*

(d) $G_1, G_2 \in \mathfrak{G}'$ *sind genau dann in* $\mathfrak{p}'$ *zueinander parallel, wenn sie es in* $\mathfrak{p}$ *sind.*

Ist $\mathfrak{p}'$ Unterstruktur von $\mathfrak{p}$, so schreiben wir dafür $\mathfrak{p}' \leqq \mathfrak{p}$. Wir sagen dann auch, $\mathfrak{p}$ sei eine *Erweiterung* oder *Oberstruktur* von $\mathfrak{p}'$, oder $\mathfrak{p}'$ sei in $\mathfrak{p}$ *eingebettet*. Übergang zu den zugehörigen p.P. ergibt offensichtlich, daß $\mathfrak{p}'^*$ Unterstruktur von $\mathfrak{p}^*$ [im Sinne von PICKERT (1955), S. 3] ist, wenn $\mathfrak{p}' \leqq \mathfrak{p}$ gilt.

Seien $\mathfrak{p}_i=(P_i, \mathfrak{G}_i)$ $[i=1, 2, \ldots]$ a.P. mit $\mathfrak{p}_1 \leqq \mathfrak{p}_2 \leqq \cdots$. Aus dieser Folge von Oberstrukturen bilden wir $\mathfrak{p}=(P, \mathfrak{G})$ auf folgende Weise: (a) Es gilt $P=\bigcup_{i=1}^{\infty} P_i$ und $\mathfrak{G}=\bigcup_{i=1}^{\infty} \mathfrak{G}_i$. (b) Sei $p \in P$ und $G \in \mathfrak{G}$, dann gibt es ein j mit $p \in P_j$ und $G \in \mathfrak{G}_j$; in $\mathfrak{p}$ soll p mit G genau dann inzidieren, wenn dies in $\mathfrak{p}_j$ der Fall ist. (c) Seien $G, G' \in \mathfrak{G}$, dann gibt es ein j mit $G, G' \in \mathfrak{G}_j$; genau dann soll G zu G' in $\mathfrak{G}$ parallel sein, wenn dies in $\mathfrak{G}_j$ der Fall ist. Man sieht sofort, daß $\mathfrak{p}$ eine a.P. ist; $\mathfrak{p}$ kann aufgefaßt werden als „kleinste" a.P., die alle $\mathfrak{p}_i$ als Unterstrukturen enthält. Wir nennen $\mathfrak{p}$ die *Vereinigung* der $\mathfrak{p}_i$ und schreiben $\mathfrak{p}=\bigcup_{i=1}^{\infty} \mathfrak{p}_i$.

§ 3. Gewebe

1. In diesem Paragraphen wollen wir eine Beziehung zwischen den Parallelstrukturen und den für viele Zweige der Geometrie wichtigen Geweben herstellen. Wir zeigen, daß die zu den Geweben dualen Strukturen spezielle p.P. sind.

DEFINITION. Es sei m eine natürliche Zahl $\geqq 3$. Ein *m-Gewebe* oder kurz Gewebe $\mathfrak{G}$ ist eine Inzidenzstruktur mit folgenden zusätzlichen Eigenschaften:

(g 1) *Jede Gerade von* $\mathfrak{G}$ *ist Hauptgerade.*

(g 2) *In* $\mathfrak{G}$ *gibt es* m *Hauptpunkte* $h_1, \ldots, h_m$; *sie liegen auf einer Geraden* H, *die keine weiteren Punkte mehr enthält.*

(g 3) $\mathfrak{G}$ *enthält mindestens zwei nicht auf* H *gelegene Punkte (sog. gewöhnliche Punkte).*

BEMERKUNG. Das hier gebrachte Axiomensystem für Gewebe ist etwas schärfer als das in PICKERT (1955), S. 42 angegebene: Statt (g 3) fordert

G. PICKERT die Existenz nur *eines* außerhalb H gelegenen Punktes [Axiom (3) in PICKERT (1955), S. 42]. Damit scheiden hier nur leicht zu übersehende Fälle aus. Etwas einschneidender ist die zweite Verschärfung. Bei PICKERT müssen nämlich die Punkte h_i $(i=1, \ldots, m)$[14]) nicht notwendig auf einer Geraden liegen. Diese Forderung ist zwar für die Untersuchung von 3-Geweben nicht wesentlich, da man bei ihnen nur die von h_i verschiedenen Punkte untersucht [vgl. PICKERT (1955), S. 42—46]. Dagegen ist sie z.B. wichtig bei 4-Geweben; bei ihnen ist es häufig zweckmäßig anzunehmen, daß die Punkte h_i auf verschiedenen Geraden liegen [PICKERT (1955), S. 61 oben]. In diesem Falle ist der jetzt zu gewinnende Zusammenhang mit den Parallelstrukturen nicht vorhanden[15]).

2. Um ihn zu finden, suchen wir die Eigenschaften der zu $\mathfrak{G}$ dualen Struktur $\mathfrak{G}'$. Den Axiomen (g 1) bis (g 3) entsprechen:

(g' 1) *Jeder Punkt von $\mathfrak{G}'$ ist Hauptpunkt.*

(g' 2) *In $\mathfrak{G}'$ gibt es m $(\geqq 3)$ Hauptgeraden $H_1, \ldots, H_m$, die durch einen Punkt h gehen; durch h geht keine weitere Gerade.*

(g' 3) *$\mathfrak{G}'$ enthält mindestens zwei nicht durch h gehende Geraden.*

Wir zeigen jetzt: *$\mathfrak{G}'$ ist eine $p.P.$*

BEWEIS. Die Gültigkeit von (p 1) und (p 2) folgt sofort aus (g' 1) bzw. (g' 2). Seien G_1 und G_2 zwei [nach (g' 3) existierende] nicht durch h gehende Geraden von $\mathfrak{G}'$. Die Punkte $G_i \cap H_k$ $(i=1, 2;\ k=2, 3)$ liegen außerhalb H_1, mindestens drei von ihnen sind verschieden und liegen nicht auf einer Geraden, das ergibt (p 3). Statt (p 4) beweisen wir weitergehend: *Jede Gerade von $\mathfrak{G}'$ enthält mindestens m Punkte.* Für eine nicht durch h gehende Gerade G sind die m Punkte $G \cap H_i$ $(i=1, \ldots, m)$ paarweise verschieden. Sei nun $h \in G$, also wegen (g' 2) o.B.d.A. $G=H_1$. Ferner sei, wegen $m \geqq 3$ ebenfalls o.B.d.A., $G_1 \cap G_2 \notin H_2$, d.h. $p=G_1 \cap H_2 \neq G_2 \cap H_2$. Dann sind die auf H_1 gelegenen m Punkte $G_2 \cap H_1$ und $\langle G_2 \cap H_k, p\rangle \cap H_1$ $(k=2, 3, \ldots, m)$ paarweise verschieden. Somit enthält H_1 und ebenso auch alle anderen Geraden durch h mindestens m Punkte.

§ 4. Kollineationen

1. Eine isomorphe Abbildung einer Inzidenzstruktur $\mathfrak{S}$ *auf sich* heißt Automorphismus oder *Kollineation* [PICKERT (1955), S. 4]. Eine Kollineation von $\mathfrak{S}$ ist also eine eineindeutige inzidenzerhaltende Abbildung der Punkte und Geraden von $\mathfrak{S}$ jeweils auf sich. Die Kollineationen von $\mathfrak{S}$ bilden eine Gruppe Γ, wenn man, wie üblich, die Hintereinanderausführung zweier Abbildungen als Gruppenverknüpfung nimmt. Das Bild eines Punktes p bzw. einer Geraden G von $\mathfrak{S}$ bei Anwendung von $\gamma \in \Gamma$ bezeichnen wir mit p^γ bzw. G^γ. Bei einer Kollineation gehen offensichtlich Hauptpunkte bzw. -geraden wieder jeweils in solche über.

[14]) In PICKERT (1955), S. 42ff. werden diese Punkte mit $A_1, \ldots, A_m$ bezeichnet.

[15]) Dies gilt z.B. auch für die in ARTZY (1953) und NAUMANN (1954, b) untersuchten 4-Gewebe.

Für eine *Kollineation γ einer a.P. $\mathfrak{p}$* verlangen wir außer der Erhaltung der Inzidenz noch die *Erhaltung der Parallelität:* Aus $G \| G'$ folgt $G^\gamma \| G'^\gamma$. Man kann γ auf die uneigentlichen Punkte von $\mathfrak{p}$ erweitern und erhält so eine Kollineation der p.P. $\mathfrak{p}^*$, welche die uneigentliche Gerade festläßt. Umgekehrt: Jede Kollineation einer p.P. $\mathfrak{P}$, die eine Hauptgerade H (als Ganzes) festläßt, induziert eine parallelenerhaltende Kollineation in der zugehörigen a.P. $\mathfrak{P}_H$. Führt man diese beiden Prozesse (in irgendeiner Reihenfolge) hintereinander aus, so kommt man zu der Kollineation zurück, von der man ausging.

2. Definition. Eine Kollineation π einer p.P. $\mathfrak{P}$, die *jeden* Punkt einer Hauptgeraden H von $\mathfrak{P}$ festläßt, nennen wir *axial*[16]) hinsichtlich H. Wir bezeichnen H als eine *Achse*[17]) von π.

Satz 4.1. *Eine nichtidentische und hinsichtlich einer Hauptgeraden H einer p.P. axiale Kollineation hat außerhalb H höchstens einen Fixpunkt.*

Beweis. Wir nehmen an, die bezüglich H axiale Kollineation π habe außerhalb H die beiden Fixpunkte p und q. Nach (p 3) gibt es mindestens einen Punkt, der weder auf H noch auf $\langle p, q\rangle$ liegt. Sei r ein solcher. Aus $p^\pi = p$ und $(\langle p, r\rangle \cap H)^\pi = \langle p, r\rangle \cap H$ folgt $\langle p, r\rangle^\pi = \langle p, r\rangle$, ebenso $\langle q, r\rangle^\pi = \langle q, r\rangle$ aus $q^\pi = q$. Wegen $r = \langle p, r\rangle \cap \langle q, r\rangle$ erhalten wir schließlich $r^\pi = r$. Mithin läßt π alle nicht auf $\langle p, q\rangle$ gelegenen Punkte fest. Aus $s \in \langle p, q\rangle$ und $s \neq p$ folgt aber $s \notin \langle p, r\rangle$, und wir können jetzt ganz entsprechend wie eben $s^\pi = s$ schließen. Alle Punkte bleiben also bei π fest. Wegen (p 4) bewegt π auch keine Gerade, mithin ist $\pi = \underline{1}$ die identische Kollineation, w.z.b.w.

Hieraus und aus Satz 2.1 folgt sofort der folgende

Zusatz. *Eine nichtidentische Kollineation einer p.P. kann axial hinsichtlich höchstens einer Geraden sein.*

Eine nichtidentische axiale Kollineation π hat also nur eine Achse, deshalb kann man in diesem Falle von *der* Achse von π reden.

Die bezüglich einer festen Achse H axialen Kollineationen bilden eine Untergruppe von Γ, die wir mit Π_H oder kurz mit Π bezeichnen, wenn kein Mißverständnis zu befürchten ist.

Eine Kollineation ζ einer p.P. $\mathfrak{P}$ heißt *zentral* hinsichtlich eines Punktes p von $\mathfrak{P}$, wenn $G^\zeta = G$ für jede durch p gehende Gerade G ist; insbesondere bleibt p selbst fest. Wir sagen, p sei ein *Zentrum* von ζ. Die hinsichtlich eines festen Punktes zentralen Kollineationen bilden eine Untergruppe von Γ.

In projektiven Ebenen gilt der Satz, daß jede axiale Kollineation auch zentral ist und umgekehrt [Baer (1942), Cor. 2.3; Pickert (1955), S. 65]. Diese einfache Beziehung braucht für allgemeine p.P. nicht mehr zu stimmen. So wird in Teil II gezeigt, daß es axiale nicht zentrale Kollineationen geben kann. Es gibt auch zentrale Kollineationen, die nicht axial sind, wie folgendes Beispiel zeigt: Sei $\mathfrak{P} = (P, \mathfrak{G})$ eine p.P. und H eine ihrer Hauptgeraden. Wir erweitern

[16]) Auch *Perspektivität* genannt.

[17]) Aus der Gültigkeit der Aussage „H ist Achse von π" soll stets folgen, daß H eine Hauptgerade ist.

$\mathfrak{P}$ zu einer Struktur $\tilde{\mathfrak{P}}=(\tilde{P}, \tilde{\mathfrak{G}})$ in folgender Weise: Zu P fügen wir zwei neue Punkte a und b hinzu, die mit H, aber keiner weiteren Geraden aus $\mathfrak{G}$ inzidieren sollen: $\tilde{P}=P\cup\{a\}\cup\{b\}$; nunmehr erweitern wir $\mathfrak{G}$ zu $\tilde{\mathfrak{G}}$ durch Hinzunahme aller Mengen $\{a, p\}$ und $\{b, p\}$ mit $p\in P-H$, d.h. wir setzen

$$\tilde{\mathfrak{G}}=\mathfrak{G}\cup\bigcup_{p\in P-H}(\{a, p\}\cup\{b, p\})$$

und fordern, daß $\{a, p\}$ nur mit a und p inzidieren soll, ebenso $\{b, p\}$ nur mit b und p. Offensichtlich ist auch $\tilde{\mathfrak{P}}$ eine p.P. Sei nun ζ eine Abbildung, die P und $\mathfrak{G}$ elementweise auf sich abbildet und für die $a^\zeta=b$ und $b^\zeta=a$, ferner $\{a, p\}^\zeta=\{b, p\}$ und $\{b, p\}^\zeta=\{a, p\}$ gilt. Man sieht sofort, daß ζ eine zentrale Kollineation von $\tilde{\mathfrak{P}}$ mit jedem von a und b verschiedenen Punkt von H als Zentrum ist. Dagegen ist ζ nicht axial. Denn H bleibt bei ζ nicht punktweise fest (a und b werden vertauscht), ist also keine Achse von ζ. Weitere Hauptgeraden enthält aber $\tilde{\mathfrak{P}}$ nicht. Sei nämlich $G\in\tilde{\mathfrak{G}}$ von H verschieden; mindestens einer der Punkte a und b liegt nicht in G; es gibt nach (p 3) mindestens einen nicht auf $G\cup H$ gelegenen Punkt p; mindestens eine der Geraden $\{a, p\}$, $\{b, p\}$ schneidet G nicht, also ist G keine Hauptgerade in $\tilde{\mathfrak{P}}$. Mithin kann ζ auch keine andere Gerade als Achse haben, denn eine Achse ist definitionsgemäß notwendig Hauptgerade.

Das Beispiel zeigt auch, daß eine nichtidentische zentrale Kollineation mehrere Zentren haben kann. Im allgemeinen gilt also das duale Gegenstück zum Zusatz von Satz 4.1 nicht. Wir können es aber unter einer kleinen Zusatzvoraussetzung beweisen:

SATZ 4.2. *Eine nichtidentische Kollineation einer p. P., in der jede Gerade mindestens drei Punkte enthält, kann zentral hinsichtlich höchstens eines Punktes sein.*

BEWEIS. Seien p und q zwei Zentren der Kollineation ζ. Ist $r\notin\langle p, q\rangle$, so ist $r^\zeta=r$ wegen $r=\langle p, r\rangle\cap\langle q, r\rangle$. Sei nun $r\in\langle p, q\rangle$. Wegen (p 3) gibt es ein $s\notin\langle p, q\rangle$. Nach unserer zusätzlichen Voraussetzung enthält die Gerade $\langle r, s\rangle$ noch einen weiteren Punkt t. Wegen $s^\zeta=s$, $t^\zeta=t$ folgt nun $r^\zeta=r$, und ζ ist die Identität.

3. Wir betrachten jetzt die axialen Kollineationen hinsichtlich einer *festen* Achse H. In diesem Falle ist es oft zweckmäßig, statt der p.P. $\mathfrak{P}$ die durch Fortnahme von H entstehende a.P. $\mathfrak{P}_H=\mathfrak{p}$ den Untersuchungen zugrunde zu legen. Eine axiale Kollineation π induziert nach Ziffer 1 in $\mathfrak{p}$ eine Kollineation, die wir ebenfalls π nennen können und wollen; offensichtlich führt sie jede Gerade in eine zu ihr parallele über: $G^\pi\| G$ für jede Gerade G von $\mathfrak{p}$. Umgekehrt erzeugt jede Kollineation von $\mathfrak{p}$ mit dieser Eigenschaft eine axiale Kollineation auf der zugehörigen p.P. $\mathfrak{p}^*=\mathfrak{P}$ (s. § 2) mit der uneigentlichen Geraden als Achse. Eine Kollineation einer a.P., die jede Gerade in eine zu ihr parallele abbildet, nennen wir *affin-axial* oder kurz *axial*[18]). Die Menge der affin-

[18]) Denn sie induziert in $\mathfrak{p}^*$ eine axiale Kollineation. Natürlich kann es in $\mathfrak{p}^*$ auch axiale Kollineationen hinsichtlich anderer Achsen geben, solche werden wir hier aber nicht untersuchen.

axialen Kollineationen bildet einen Normalteiler Π in der Gruppe Γ aller Kollineationen von $\mathfrak{p}$, denn es ist $G^{\gamma\pi\gamma^{-1}} \| G^{\gamma\gamma^{-1}} = G$ für $\gamma \in \Gamma$, $\pi \in \Pi$.

Satz 4.1 besagt affin, daß eine nichtidentische affin-axiale Kollineation höchstens einen (eigentlichen) Fixpunkt haben kann. Eine affin-axiale Kollineation von $\mathfrak{p}$ die entweder alle oder keinen Punkt festläßt, ist eine *Translation*, eine, die alle oder genau einen Punkt festhält, heißt *Streckung*. Auch die von ihnen in der zugehörigen p. P. $\mathfrak{p}^*$ induzierten Kollineationen bezeichnen wir als Translationen bzw. Streckungen.

Die Identität ist die einzige Kollineation, die sowohl Translation als auch Streckung ist. Jede affin-axiale Kollineation ist eine Translation oder eine Streckung. Jede Streckung ist zentral bezüglich ihres Fixpunktes. Dagegen braucht, im Gegensatz zu Ebenen, nicht jede Translation *zentral* zu sein (also ein Zentrum zu besitzen), wie in Teil II gezeigt wird. Falls eine Translation $\tau \neq \underline{1}$ ein Zentrum z besitzt, so ist dies ein uneigentlicher Punkt der a. P. $\mathfrak{p}$, denn im Eigentlichen hat τ definitionsgemäß keinen Fixpunkt, also auch kein Zentrum. Genau die durch z gehenden Geraden von $\mathfrak{p}$ bleiben bei Anwendung von τ (als Ganzes) fest, mit einer Geraden also auch die zu ihr parallelen und nur diese. Sei umgekehrt $\tau \neq \underline{1}$ eine Translation mit folgenden Eigenschaften: (a) Es gibt eine Gerade G mit $G^\tau = G$, (b) mit $G \| G'$ ist $G'^\tau = G'$ gleichwertig. Dann ist τ zentral mit dem uneigentlichen Punkt von G als Zentrum. Man kann eine zentrale Translation $\tau \neq \underline{1}$ auch durch die Gültigkeit von $\langle p, p^\tau \rangle \| \langle q, q^\tau \rangle$ für alle Punkte p, q kennzeichnen.

4. Die Streckungen mit dem Fixpunkt p bilden eine Untergruppe Π_p der Gruppe Π aller affin-axialen Kollineationen; die Elemente von Π_p nennen wir auch *p-Streckungen*. Nach Satz 4.1 haben Π_p und Π_q für $p \neq q$ nur die Identität gemeinsam: $\Pi_p \cap \Pi_q = \{\underline{1}\}$. Es ist allgemein nicht bekannt, ob die Translationen eine Untergruppe von Π bilden. Dies gilt zwar für Ebenen [Baer (1942), Cor. 2.4; Pickert (1955), S. 65], der Beweis läßt sich jedoch auf den Fall beliebiger Parallelstrukturen nicht übertragen, da in ihm wesentlich benutzt wird, daß zwei nichtparallele Geraden sich schneiden. Trivialerweise gilt aber jedenfalls: Mit τ sind auch τ^{-1} und $\pi\tau\pi^{-1}$ ($\pi \in \Pi$) Translationen; dasselbe gilt übrigens auch für zentrale Translationen. *Bilden also die Translationen (zentralen Translationen) eine Untergruppe, so auch einen Normalteiler.* Ferner kann man zeigen:

Satz 4.3. *Die Translationen einer a. P. bilden genau dann eine Gruppe, wenn jede Translation durch ihre Wirkung auf einen Punkt eindeutig bestimmt ist.*

Beweis. Es sei T die Menge aller Translationen der a. P. $\mathfrak{p}$.

(a) Sei T eine Gruppe; aus $p^\tau = p^{\tau'}$ für einen Punkt p von $\mathfrak{p}$ folgt $p^{\tau\tau'^{-1}} = p$ mit $\tau\tau'^{-1} \in \mathsf{T}$, also $\tau\tau'^{-1} = \underline{1}$, $\tau = \tau'$.

(b) Sei T keine Gruppe, dann gibt es $\tau, \tau' \in \mathsf{T}$ mit $\tau\tau'^{-1} \notin \mathsf{T}$, d. h. $\tau\tau'^{-1}$ ist eine nichtidentische Streckung mit einem Zentrum p. Einerseits ist $\tau \neq \tau'$, andererseits aber $p^{\tau\tau'^{-1}} = p$, mithin $p^\tau = p^{\tau'}$; es gibt somit zwei verschiedene Translationen, die beide dieselbe Wirkung auf p haben.

SATZ 4.4. *Jede zentrale Translation ist durch ihre Wirkung auf einen Punkt eindeutig bestimmt.*

BEWEIS. Sei $\tau \neq \underline{1}$ eine zentrale Translation von $\mathfrak{p}$, die p in $p' \neq p$ überführt. Nach (a 3) [vgl. § 2] gibt es in $\mathfrak{p}$ mindestens einen nicht auf $\langle p, p' \rangle$ gelegenen Punkt q. Es ist $\langle p, q \rangle \| \langle p', q^\tau \rangle$, weil τ eine Translation, und $\langle p, p' \rangle \| \langle q, q^\tau \rangle$, weil τ zentral ist. Aus $q^\tau = \langle p', q^\tau \rangle \cap \langle q, q^\tau \rangle$ und der Eindeutigkeit des Ziehens von Parallelen [Axiom (a 2)] folgt, daß q^τ eindeutig bestimmt ist. Durch die Wirkung auf p ist also τ eindeutig auf den Punkten außerhalb von $\langle p, p' \rangle$ festgelegt. Ist aber $r \in \langle p, p' \rangle$, so $r \notin \langle q, q^\tau \rangle$, und wir können nun genau so die Eindeutigkeit von r^τ zeigen. Damit ist der Satz vollständig bewiesen.

Seien τ, τ' zwei zentrale Translationen. Wäre $\tau\tau'$ eine p-Streckung, so $p^{\tau\tau'} = p$, $p^\tau = p^{\tau'^{-1}}$, nach dem eben bewiesenen Satz also $\tau\tau' = \underline{1}$. Mithin gilt der

ZUSATZ. *Das Produkt zweier zentraler Translationen ist eine Translation (von der nicht bekannt ist, ob sie notwendig zentral ist).*

Für spätere Untersuchungen ist folgender Satz nützlich.

SATZ 4.5. *Es seien τ und τ' zwei zentrale Translationen mit verschiedenen Zentren; dann ist $\tau\tau' = \tau'\tau$.*

Wir führen den Beweis dieses Satzes ganz entsprechend wie bei Ebenen [vgl. z.B. PICKERT (1955), S. 199]: Wir können o.B.d.A. $\tau, \tau' \neq \underline{1}$ annehmen. Weil τ und τ' verschiedene Zentren haben sollen, gilt $\langle p, p^\tau \rangle \neq \langle p, p^{\tau'} \rangle$ für jeden Punkt p. Da τ bzw. τ' zentrale Translationen sind, folgt $\langle p, p^{\tau'} \rangle \| \langle p, p^{\tau'} \rangle^\tau = \langle p^\tau, p^{\tau'\tau} \rangle$ und $\langle p, p^\tau \rangle \| \langle p^{\tau'}, p^{\tau'\tau} \rangle$ bzw. $\langle p, p^\tau \rangle \| \langle p^{\tau'}, p^{\tau\tau'} \rangle$ und $\langle p, p^{\tau'} \rangle \| \langle p^\tau, p^{\tau\tau'} \rangle$, hieraus $p^{\tau\tau'} = \langle p^\tau, p^{\tau\tau'} \rangle \cap \langle p^{\tau'}, p^{\tau\tau'} \rangle = \langle p^\tau, p^{\tau'\tau} \rangle \cap \langle p^{\tau'}, p^{\tau'\tau} \rangle = p^{\tau'\tau}$, also $\tau\tau' = \tau'\tau$.

Für manche Spezialfälle ist es möglich, zu zeigen, daß die Translationen eine Gruppe bilden.

SATZ 4.6. *Jede der folgenden Bedingungen für eine a.P. $\mathfrak{p}$ ist hinreichend dafür, daß deren Translationen eine Gruppe bilden:*

(a) *$\mathfrak{p}$ ist endlich,*

(b) *jede Translation von $\mathfrak{p}$ ist zentral,*

(c) *$\mathfrak{p}$ ist eine affine Ebene.*

BEWEIS. Daß (a) hinreicht, folgt leicht aus einem Satz von FROBENIUS über endliche Permutationsgruppen, s. hierzu Satz 5.2 des nächsten Paragraphen. Daß (b) genügt, ergibt sich aus den Sätzen 4.3 und 4.4. Schließlich folgt (c) aus (b), da in einer Ebene jede Translation zentral ist.

5. Sei die a.P. $\hat{\mathfrak{p}} = (\hat{P}, \hat{\mathfrak{G}})$ eine Unterstruktur der a.P. $\mathfrak{p} = (P, \mathfrak{G})$ [vgl. § 2, Ziffer 5]. Jede Kollineation γ von $\mathfrak{p}$, für die $\hat{P}^\gamma = \hat{P}$ und $\hat{\mathfrak{G}}^\gamma = \hat{\mathfrak{G}}$ gilt [19]), induziert auf $\hat{\mathfrak{p}}$ eine Kollineation $\hat{\gamma}$, die durch $p^{\hat{\gamma}} = p^\gamma$ und $G^{\hat{\gamma}} = G^\gamma$ für $p \in \hat{P}$ bzw. $G \in \hat{\mathfrak{G}}$ definiert ist. Die Menge dieser so festgelegten Kollineationen bildet

[19]) Enthält jede Gerade aus $\hat{\mathfrak{G}}$ mindestens zwei Punkte aus $\hat{P}$, so genügt es offenbar, $\hat{P}^\gamma = \hat{P}$ zu fordern.

eine Gruppe $\hat{\Gamma}$, die vermöge der Abbildung $\gamma \to \hat{\gamma}$ homomorphes Bild der Kollineationsgruppe Γ von $\mathfrak{p}$ ist.

DEFINITION. Eine Unterstruktur $\hat{\mathfrak{p}} = (\hat{P}, \hat{\mathfrak{G}})$ der a.P. $\mathfrak{p} = (P, \mathfrak{G})$ heißt *kollineationstreu* in $\mathfrak{p}$, wenn gilt:

(a) *Jede Kollineation γ von $\mathfrak{p}$ erfüllt $\hat{P}^\gamma = \hat{P}$ und $\hat{\mathfrak{G}}^\gamma = \hat{\mathfrak{G}}$.*

(b) *Zu jeder Kollineation γ' von $\hat{\mathfrak{p}}$ gibt es genau eine Kollineation γ von $\mathfrak{p}$ derart, daß die von ihr auf $\hat{\mathfrak{p}}$ induzierte Kollineation $\hat{\gamma}$ mit γ' übereinstimmt.*

BEMERKUNG. Aussage (a) besagt, daß jede Kollineation von $\mathfrak{p}$ eine Kollineation auf $\hat{\mathfrak{p}}$ induziert; (b) bedeutet, daß man umgekehrt jede Kollineation von $\hat{\mathfrak{p}}$ auf diese Weise aus genau einer Kollineation von $\mathfrak{p}$ gewinnen kann. Liegt $\hat{\mathfrak{p}}$ kollineationstreu in $\mathfrak{p}$, so stimmen die Kollineationsgruppen von $\mathfrak{p}$ und $\hat{\mathfrak{p}}$ überein, wenn wir (was in diesem Falle möglich ist) γ mit $\hat{\gamma}$ identifizieren.

Sei jetzt insbesondere $\gamma = \tau$ eine Translation von $\mathfrak{p}$ mit $\hat{P}^\tau = \hat{P}$ und $\hat{\mathfrak{G}}^\tau = \hat{\mathfrak{G}}$. Dann ist offensichtlich $\hat{\tau}$ eine Translation auf $\hat{\mathfrak{p}}$; ist τ zentral auf $\mathfrak{p}$, so auch $\hat{\tau}$ auf $\hat{\mathfrak{p}}$. Um die von τ auf $\hat{\mathfrak{p}}$ induzierte Translation $\hat{\tau}$ bilden zu können, genügt es, $\hat{P}^\tau = \hat{P}$ zu fordern. Aus $p \in G \in \hat{\mathfrak{G}}$, $p \in \hat{P}$ folgt nämlich wegen $G^\tau = \langle G \| p^\tau \rangle$ (s. § 2, Ende von Ziffer 2) und der Gültigkeit von (a 2) in $\hat{\mathfrak{p}}$ auch $G^\tau \in \hat{\mathfrak{G}}$, also $\hat{\mathfrak{G}}^\tau = \hat{\mathfrak{G}}$. Aus $\hat{\tau} = \underline{1}$ folgt, daß $\hat{\tau}$ und damit auch τ mindestens zwei Punkte von $\mathfrak{p}$ festläßt; nach Satz 4.1 ist also auch $\tau = \underline{1}$. Wegen der Homomorphieeigenschaft der Zuordnung $\tau \to \hat{\tau}$ folgt nunmehr $\tau_1 = \tau_2$ aus $\hat{\tau}_1 = \hat{\tau}_2$ für zwei Translationen τ_1 und τ_2.

Analog wie oben „kollineationstreu" können wir jetzt den Begriff „translationstreu" einführen. Eine Unterstruktur $\hat{\mathfrak{p}}$ der a.P. $\mathfrak{p}$ heißt *translationstreu* in $\mathfrak{p}$, wenn folgende beiden Bedingungen erfüllt sind:

(a') *Für jede Translation τ von $\mathfrak{p}$ gilt $\hat{P}^\tau = \hat{P}$.*

(b') *Zu jeder Translation τ' von $\hat{\mathfrak{p}}$ gibt es eine Translation τ von $\mathfrak{p}$ derart, daß die von ihr auf $\hat{\mathfrak{p}}$ induzierte Kollineation $\hat{\tau} = \tau'$ ist.*

Wegen der beiden eben gemachten Bemerkungen folgt $\hat{\mathfrak{G}}^\tau = \hat{\mathfrak{G}}$ aus (a') und die Existenz nur einer Translation τ mit $\hat{\tau} = \tau'$ aus (b'). Damit ist die Analogie zwischen den Begriffen „kollineationstreu', und „translationstreu" vollständig hergestellt. Die Translationsgruppen von $\mathfrak{p}$ und $\hat{\mathfrak{p}}$ stimmen unter der Annahme der Translationstreue überein bei Identifizierung von τ mit $\hat{\tau}$.

Offensichtlich gilt die folgende Transitivitätsbeziehung:

SATZ 4.7. *Ist $\mathfrak{p}_1$ kollineationstreu (translationstreu) in $\mathfrak{p}_0$ und $\mathfrak{p}_2$ kollineationstreu (translationstreu) in $\mathfrak{p}_1$, so ist $\mathfrak{p}_2$ kollineationstreu (translationstreu) in $\mathfrak{p}_0$.*

Für die Aussage, die Unterstruktur $\mathfrak{p}$ von $\tilde{\mathfrak{p}}$ liege in $\tilde{\mathfrak{p}}$ translationstreu, sagen wir auch, $\tilde{\mathfrak{p}}$ sei eine *translationstreue Erweiterung* von $\mathfrak{p}$, oder $\tilde{\mathfrak{p}}$ sei translationstreu über[20]) $\mathfrak{p}$.

Beispiele translationstreuer Erweiterungen einer a.P. $\mathfrak{p} = (P, \mathfrak{G})$ erhält man folgendermaßen: Sei J eine Indexmenge beliebiger Mächtigkeit; dann werde die Erweiterung $\mathfrak{p}^{(J)} = (P, \mathfrak{G}^{(J)})$ von $\mathfrak{p}$ auf folgende Weise definiert:

[20]) Eine entsprechende Terminologie soll auch für kollineationstreu gelten.

Die Punkte von $\mathfrak{p}^{(J)}$ stimmen mit denen von $\mathfrak{p}$ überein. Die Geradenmenge $\mathfrak{G}^{(J)}$ von $\mathfrak{p}^{(J)}$ bestehe aus den Geraden von $\mathfrak{G}$ und den Paaren (p, j) mit $p \in P$ und $j \in J$:

$$\mathfrak{G}^{(J)} = \mathfrak{G} \cup \{(p, j) \mid p \in P,\ j \in J\}.$$

Die Geraden (p, j) sollen mit p, aber keinem anderen Punkt inzidieren. Zwei Geraden (p, j) und (p', j') $(p, p' \in P;\ j, j' \in J)$ seien genau dann zueinander parallel, wenn $j = j'$ ist. Keine der Geraden aus $\mathfrak{G}$ sei zu einem (p, j) parallel, während sich die in $\mathfrak{p}$ gegebene Parallelität auf $\mathfrak{p}^{(J)}$ überträgt. Man sieht nun, daß $\mathfrak{p}^{(J)}$ eine a.P. und Erweiterung von $\mathfrak{p}$ ist. Jede Translation τ von $\mathfrak{p}$ induziert in $\mathfrak{p}^{(J)}$ eine Translation (die wir ebenfalls τ nennen können), indem wir $(p, j)^\tau = (p^\tau, j)$ festsetzen, und umgekehrt erhält man so alle Translationen von $\mathfrak{p}^{(J)}$. Nun folgt sofort, daß $\mathfrak{p}^{(J)}$ eine translationstreue Erweiterung[21]) von $\mathfrak{p}$ ist.

In den Teilen III und IV werden wir weitere Beispiele translationstreuer Erweiterungen kennenlernen[22]).

Schließlich sei noch vermerkt, daß ganz analoge Betrachtungen für die affin-axialen Kollineationen (anstelle der allgemeinen Kollineationen oder Translationen) durchgeführt werden können.

§ 5. Sätze über spezielle Permutationsgruppen

1. In diesem Paragraphen beweisen wir einige Sätze über gewisse spezielle Permutationsgruppen, die für die Untersuchung endlicher a.P. von Interesse sind; insbesondere gelingt es uns mit ihnen, Satz 4.6 (a) zu beweisen. Es sei P eine *endliche* Menge von Elementen $p, q, \ldots$, die wir *Punkte* nennen wollen; ferner sei Π eine Gruppe von Permutationen der Punkte von P mit folgender zusätzlicher Eigenschaft:

(F) *Aus* $p, q \in P$, $p \neq q$, $\pi \in \Pi$ *und* $p^\pi = p$, $q^\pi = q$ *folgt* $\pi = 1$, *d.h. jede Permutation von* Π *ist durch ihre Wirkung auf zwei Punkte von* P *eindeutig bestimmt.*

Permutationsgruppen mit dieser Eigenschaft nennen wir *Frobenius-Gruppen*[23]). Aus Satz 4.1 folgt, daß z.B. die Gruppe der affin-axialen Kollineationen einer (endlichen) a.P. eine Frobenius-Gruppe ist, wenn wir sie als Permutationsgruppe auf ihren Punkten auffassen. Die Menge Π_p aller $\pi \in \Pi$ mit $p^\pi = p$ für ein $p \in P$ ist eine Untergruppe von Π, die Gruppe der *p-Streckungen*. — Eine Permutation von Π, die alles oder keinen Punkt von P festläßt, nennen wir *Translation*; die Menge aller Translationen von Π bezeichnen wir mit T.

Es ist das Ziel dieses Paragraphen, zu zeigen, daß T eine Untergruppe (und damit Normalteiler) in Π ist. Für ein auf P transitives Π besagt dies ein bekannter Satz von FROBENIUS, den man bisher allerdings nur mit Hilfe

[21]) Ist insbesondere $J = \emptyset$, so stimmt $\mathfrak{p}^{(J)}$ mit $\mathfrak{p}$ überein.

[22]) Während bei den hier betrachteten Beispielen die Punktmenge ungeändert blieb, werden dort sowohl die Punkt- als auch die Geradenmenge echt erweitert.

[23]) Diese Definition weicht von der in WIELANDT (1955), S. 10 gegebenen etwas ab.

der Theorie der Gruppencharaktere beweisen kann [s. hierzu etwa SPEISER (1937), S. 202]:

SATZ 5.1 (FROBENIUS). *Die Translationen einer transitiven Frobenius-Gruppe bilden einen transitiven Normalteiler.*

Wir nehmen jetzt an, daß P durch Π in n (>1) Transitivitätsgebiete P_i ($i=1, \ldots, n$) zerfalle[24]: $P=\bigcup_{i=1}^{n} P_i$. Die Gruppe Π^{P_i} der von Π auf P_i induzierten Permutationen (der *Konstituent* von Π auf P_i) ist eine auf P_i transitive Frobenius-Gruppe. Es sei T_i die Menge derjenigen Permutationen von Π, die auf P_i Translationen induzieren; T_i ist eine Untergruppe von Π, denn wenn $\tau, \tau' \in \mathsf{T}_i$ auf P_i Translationen induzieren, so auch $\tau\tau'^{-1}$ nach Satz 5.1, angewandt auf Π^{P_i}. Mithin ist auch $\mathsf{T}^*=\bigcap_{i=1}^{n} \mathsf{T}_i$ eine Untergruppe von Π.

Enthält ein Transitivitätsgebiet von Π nur einen Punkt p, so ist $\Pi=\Pi_p$, also $\mathsf{T}=\{1\}$ eine Gruppe. Wir nehmen jetzt an, daß jedes Transitivitätsgebiet P_i mindestens zwei Punkte enthalte: $|P_i|>1$[25]. In diesem Fall ist $\mathsf{T}=\mathsf{T}^*$. Jede Translation von P induziert nämlich eine Translation in P_i, das bedeutet $\mathsf{T}\subseteq\mathsf{T}^*$. Sei umgekehrt $\tau\in\mathsf{T}^*$. Nach Definition läßt τ in jedem P_i alles oder keinen Punkt fest. Läßt τ in einem P_i alle Punkte fest, wegen $|P_i|>1$ also mindestens zwei, so ist $\tau=1$, weil Π, also auch T^*, eine Frobenius-Gruppe ist. Somit ist τ eine Translation, $\mathsf{T}^*\subseteq\mathsf{T}$, also $\mathsf{T}^*=\mathsf{T}$, was behauptet war. Wir erhalten also [s. auch FROBENIUS (1902), oder WIELANDT (1955), S. 11 oben]:

SATZ 5.2. *Die Translationen einer Frobenius-Gruppe bilden einen Normalteiler.*

In diesem Satz ist Satz 4.6 (a) als Spezialfall enthalten.

2. Über die Streckungen einer Frobenius-Gruppe gibt Auskunft der

SATZ 5.3 [FROBENIUS (1902)]. *Sei Π eine Frobenius-Gruppe auf P. Die Menge aller Punkte $p\in P$ mit $\Pi_p>\{1\}$ ist entweder leer oder stimmt mit genau einem Transitivitätsgebiet P_1 von Π überein. In diesem Falle gilt:*

(a) *Für alle $p\in P_1$ ist $\Pi_{p^\pi}=\pi^{-1}\Pi_p\pi$ ($\pi\in\Pi$) und $\Pi=\Pi_p\mathsf{T}$.*

(b) *P_1 ist auch Transitivitätsgebiet bezüglich der Gruppe T der Translationen von Π. Insbesondere ist also Π genau dann transitiv, wenn T es ist.*

(c) *Die übrigen Transitivitätsgebiete von Π zerfallen in ebensoviele Transitivitätsgebiete hinsichtlich T wie Π_p ($p\in P_1$) Elemente hat. Jedes Transitivitätsgebiet T von T ist ein Block* [s. WIELANDT (1955), § 6] *von Π_p, d.h. aus $\pi\in\Pi_p$ folgt $T^\pi=T$ oder $T^\pi\cap T=\emptyset$.*

Zum Beweis verweisen wir auf die entsprechenden Untersuchungen über Perspektivitäten endlicher Ebenen in ANDRÉ (1955); die dort angegebenen Beweise kann man direkt übertragen, da in ihnen lediglich benutzt wurde,

[24] Ist p ein fester Punkt von P_i, so besteht also P_i aus genau allen p^π mit $\pi\in\Pi$, d.h. es ist $P_i=p^\Pi$.

[25] Mit $|M|$ bezeichnen wir die Mächtigkeit einer Menge M.

daß die Perspektivitäten (affin-axialen Kollineationen) eine Frobenius-Gruppe bilden. Lediglich die Blockeigenschaft der Transitivitätsgebiete von T hinsichtlich Π_p wurde dort nicht bewiesen; sie folgt aber leicht so [WIELANDT (1955), § 6]: Ist $q \in T^\pi \cap T$ für ein Transitivitätsgebiet T von T, ferner $r \in T$, $s \in T^\pi$, so gibt es $\tau, \tau' \in \mathsf{T}$ mit $r = p^\tau$ und $s = q^{\tau'}$ (denn mit T ist auch T^π Transitivitätsgebiet von T), mithin liegen r und s in demselben Transitivitätsgebiet von T, also gilt $T = T^\pi$, was behauptet war.

Literatur

ANDRÉ, J.: Über Perspektivitäten in endlichen projektiven Ebenen. Arch. Math. **6**, 29–32 (1955).

— Über Perspektivitäten in endlichen affinen Ebenen. Arch. Math. **9**, 228–235 (1958).

— Affine Ebenen mit genügend vielen Translationen. Math. Nachr. **19**, 204–210 (1959).

ARTZY, R.: Eigenschaften von ebenen Viergeweben allgemeiner Lage. Math. Ann. **126**, 336–342 (1953).

BAER, R.: Homogeneity of projective planes. Amer. J. Math. **64**, 137–152 (1942).

BLASCHKE, W., u. G. BOL: Geometrie der Gewebe. Berlin-Göttingen-Heidelberg: Springer 1938.

FROBENIUS, G.: Über primitive Gruppen des Grades n und der Klasse $n-1$. S.-B. Akad. Berlin 455–459 (1902).

HALL, M.: Projective planes. Trans. Amer. Math. Soc. **54**, 229–277 (1943).

HUGHES, D. R.: A class of non-desarguesian projective planes. Canad. J. Math. **9**, 378–388 (1957).

LOMBARDO-RADICE, L.: I piani di refrazioni. Rend. Mat. Roma **13**, 130–139 (1954).

MOULTON, F. R.: A simple non-desarguesian plane geometry. Trans. Amer. Math. Soc. **3**, 192–195 (1902).

NAUMANN, H.: Stufen der Begründung der ebenen affinen Geometrie. Math. Z. **60**, 120–141 (1954, a).

— Über das zweite Distributivgesetz im Zusammenhang mit den Geweben von Herrn ARTZY. Math. Ann. **128**, 92–94 (1954, b).

PICKERT, G.: Projektive Ebenen. Berlin-Göttingen-Heidelberg: Springer 1955.

REIDEMEISTER, K.: Vorlesungen über Grundlagen der Geometrie. Berlin-Göttingen-Heidelberg: Springer 1930.

SALZMANN, H.: Kompakte zweidimensionale Ebenen. Arch. Math. **9**, 447–454 (1958).

SPEISER, A.: Theorie der Gruppen von endlicher Ordnung. Berlin-Göttingen-Heidelberg: Springer 1937.

WIELANDT, H.: Permutationsgruppen. Hektographierte Vorlesungsausarbeitung, Universität Tübingen 1955.

ZAPPA, G.: Piani affini con traslationi. Ric. Matematica **7**, 1–13 (1959).

Braunschweig, Mathematisches Institut der Technischen Hochschule

(Eingegangen am 13. Oktober 1960)

VOGEL, W.
Math. Zeitschr. 76, 103–115 (1961)

Lineare Programme und allgemeine Vertretersysteme

Von

WALTER VOGEL

1. Einleitung. Unter einem linearen Programm versteht man die Bestimmung des Extremwertes einer linearen Funktion (in mehreren Veränderlichen) unter linearen Nebenbedingungen. In den Jahren 1946/47 fand G. B. DANTZIG dafür ein Rechenverfahren (die sog. *Simplex-Methode*), und seit dieser Zeit hat man eine große Zahl von Anwendungsmöglichkeiten auf wirtschaftlichem und militärischem Bereich gefunden. In den letzten Jahren ist die Theorie der linearen Programme aber auch zur Lösung kombinatorischer Probleme herangezogen worden. Auf diese Weise haben z. B. FORD, FULKERSON, HOFFMAN und KUHN in [4] und [11] notwendige und hinreichende Bedingungen für die Existenz von Vertretersystemen gefunden. Ein Vertretersystem für ein System von Teilmengen $\mathscr{A}_i$ einer Menge $\mathscr{A}$ ist eine Teilmenge $\mathscr{S}$ von $\mathscr{A}$ mit der Eigenschaft, daß für jedes $\mathscr{A}_i$ ein Element a aus $\mathscr{S}$ mit $a \in \mathscr{A}_i$ gefunden werden kann. In [11] wird zusätzlich verlangt, daß einige vorgeschriebene Elemente von $\mathscr{A}$ zu $\mathscr{S}$ gehören sollen und daß die Vertreter verschiedener Teilmengen voneinander verschieden sein sollen. In [4] wird der Fall behandelt, daß das Element a_k wenigstens α_k-mal und höchstens β_k-mal als Vertreter einer Menge auftritt.

Wir behandeln hier Vertretersysteme mit anderen Nebenbedingungen. Ihre Definition findet man im Abschnitt 4. Als Vorbereitung wollen wir in diesem und dem nächsten Abschnitt einiges über lineare Programme zusammenstellen. In Abschnitt 3 werden dann Sätze über Matrizen aus Nullen und Einsen bewiesen. Diese Sätze liefern fast unmittelbar die Aussagen über Vertretersysteme aus dem Abschnitt 4.

Wir benutzen folgende Bezeichnungsweisen: a, b, x, y usw. stellen Spalten-Vektoren und A, B usw. stellen Matrizen dar. a' (bzw. A') bedeutet den Übergang zum transponierten Vektor (bzw. zur transponierten Matrix). So ist z. B. $x' = (\xi_1, \xi_2, \ldots, \xi_n)$. Das Zeichen 0 schreiben wir sowohl für die Ziffer Null als auch für den Null-Vektor oder die Null-Matrix. Die Dimension wird bei Matrizen und Vektoren nicht eigens angegeben, sie kann stets aus dem Zusammenhang entnommen werden. Wenn z. B. A eine Matrix mit m Zeilen und n Spalten ist und wenn die Gleichung $Ax = b$ angeschrieben wird, so ist vorausgesetzt, daß x ein n-dimensionaler und b ein m-dimensionaler Spaltenvektor ist. I bedeutet immer die Einheitsmatrix. Wir setzen Matrizen und Vektoren zu neuen Matrizen zusammen; z. B. $A = (a_1, a_2, \ldots, a_n)$ oder $B = (A, I)$. Die Formel $x \geqq 0$ bedeutet: jede Komponente von x ist größer oder gleich Null.

Mit diesen Bezeichnungen schreiben wir nun ein lineares Programm an.

$$\mathfrak{P}\text{: Suche } \max\left\{c'x+d'y \;\middle|\; \begin{pmatrix}x\\y\end{pmatrix}\in\mathscr{P}\right\}.$$

Dabei sei

$$\mathscr{P}=\left\{\begin{pmatrix}x\\y\end{pmatrix}\;\middle|\; \begin{matrix}A_1x+A_2y\leqq a\\ B_1x+B_2y=b\end{matrix},\ x\geqq 0\right\}.$$

Es hat sich folgende Terminologie eingebürgert:

Jeder Vektor $\begin{pmatrix}x\\y\end{pmatrix}$ mit $\begin{pmatrix}x\\y\end{pmatrix}\in\mathscr{P}$ heißt eine zulässige Lösung.

Jede zulässige Lösung, welche das Maximum liefert, heißt eine optimale Lösung.

Neben $\mathfrak{P}$ wollen wir das Programm $\mathfrak{Q}$ einführen.

$$\mathfrak{Q}\text{: Suche } \min\left\{a'u+b'v \;\middle|\; \begin{pmatrix}u\\v\end{pmatrix}\in\mathscr{Q}\right\}.$$

Dabei sei

$$\mathscr{Q}=\left\{\begin{pmatrix}u\\v\end{pmatrix}\;\middle|\; \begin{matrix}A_1'u+B_1'v\geqq c\\ A_2'u+B_2'v=d\end{matrix},\ u\geqq 0\right\}.$$

Die Programme $\mathfrak{P}$ und $\mathfrak{Q}$ heißen zueinander dual. Ihre gegenseitige Beziehung wird in Satz 1.1 ausgesprochen.

SATZ 1.1. *Die drei folgenden Aussagen sind gleichwertig.*

(i) *Das Programm* $\mathfrak{P}$ *hat eine optimale Lösung.*

(ii) *Das Programm* $\mathfrak{Q}$ *hat eine optimale Lösung.*

(iii) $\mathfrak{P}$ *und* $\mathfrak{Q}$ *besitzen zulässige Lösungen.*

Gilt eine der drei Aussagen (und damit auch die beiden anderen), so ist

$$\max(c'x+d'y)=\min(a'u+b'v).$$

Dieser Satz heißt der *Dualitätssatz* für lineare Programme. Beweise findet man z.B. in [5] oder [2].

2. Unimodulare Matrizen und lineare Programme. Der Inhalt dieses Abschnitts kann folgendermaßen wiedergegeben werden: Eine unimodulare Matrix kann zur Definition einer konvexen Menge mit ganzzahligen Ecken (d.h. mit Ecken, deren Koordinaten ganze Zahlen sind) benutzt werden. Wenn ein lineares Programm eine optimale Lösung besitzt, so besitzt es auch eine optimale Ecke. Mit Hilfe unimodularer Matrizen kann man daher lineare Programme mit ganzzahligen Lösungen konstruieren. Dies ist im wesentlichen schon bekannt, doch sind die folgenden Ausführungen als Einleitung für die nächsten Abschnitte notwendig.

DEFINITION. Eine Matrix heißt *unimodular*, wenn jede ihrer Unterdeterminanten einen der Werte $+1$, 0 oder -1 hat.

Man sieht leicht: Jede Teilmatrix einer unimodularen Matrix ist wieder unimodular. Mit A ist auch A' unimodular. Als Elemente einer unimodularen Matrix kommen nur die Zahlen $+1$, 0 oder -1 in Betracht, denn die Elemente sind spezielle Unterdeterminanten.

Die Matrix des weiter unten stehenden Gleichungssystemes (in $x_{i,k}$) ist unimodular. Dies folgt sofort aus Theorem 3 in [9, S. 233] oder aus [8, S. 252].

$$\sum_{i=1}^{m} x_{i,k} = r_k, \qquad k = 1, 2, \ldots, n,$$

$$\sum_{k=1}^{n} x_{i,k} = s_i, \qquad i = 1, 2, \ldots, m.$$

Wir schreiben als Beispiel eine Matrix für $m=4$ und $n=3$ an.

$$\begin{pmatrix} 1&1&1&1&&&&&&&&\\ &&&&1&1&1&1&&&&\\ &&&&&&&&1&1&1&1\\ 1&&&&1&&&&1&&&\\ &1&&&&1&&&&1&&\\ &&1&&&&1&&&&1&\\ &&&1&&&&1&&&&1 \end{pmatrix}$$

Für den folgenden Satz erinnern wir zunächst an die Definition der Ecke einer konvexen Menge.

DEFINITION. x heißt eine *Ecke* der konvexen Menge $\mathfrak{M}$, wenn aus $x, y, z \in \mathfrak{M}$ und $x = (y+z)/2$ stets folgt, daß $x = y = z$ ist.

SATZ 2.1. *Sei A unimodular und seien b und c ganzzahlige Vektoren. Dann besitzen die konvexen Mengen*

$$\mathfrak{M}_1 = \{x \mid A\,x \leqq b,\ 0 \leqq x\}$$

und

$$\mathfrak{M}_2 = \{x \mid A\,x \leqq b,\ 0 \leqq x \leqq c\}$$

nur ganzzahlige Ecken.

Natürlich kann $\mathfrak{M}_i$ $(i=1, 2)$ auch leer sein. Wenn dies aber nicht der Fall ist, so gibt es wenigstens eine Ecke in $\mathfrak{M}_i$. Den Beweis des Satzes findet man in [9, Theorem 2, S. 225].

Wir wollen nun zeigen:

LEMMA 2.2. *Wenn $A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$ unimodular ist, so sind auch $B = \begin{pmatrix} A_1, I \\ A_2, 0 \end{pmatrix}$ und $C = (A', -A_2')$ unimodular.*

BEWEIS. 1. Man entwickelt jede Unterdeterminante von B zuerst nach den Spalten, welche aus dem Teil $\begin{pmatrix} I \\ 0 \end{pmatrix}$ stammen. Auf diese Weise erhält man $+1$, -1 oder 0 mal Wert einer Unterdeterminante von A. 2. Eine Unterdeterminante von C, welche zwei entsprechende Spalten aus A' und A_2' enthält, hat den Wert Null. Jede andere Unterdeterminante kann durch

Multiplikation der Spalten aus $-A_2'$ mit -1 und anschließende Permutation der Spalten in eine Unterdeterminante von A' übergeführt werden. Da sie dabei nur den Faktor $+1$ oder -1 erhält, so folgt die Behauptung des Lemmas.

Schreibt man $A\,x=b$ in der Form $\begin{pmatrix} A \\ -A \end{pmatrix} x \leqq \begin{pmatrix} b \\ -b \end{pmatrix}$, so liefert der Satz 2.1 unmittelbar den

SATZ 2.3. *Sei A unimodular und seien b und c ganzzahlige Vektoren. Dann haben die konvexen Mengen*

$$\mathfrak{X}_3 = \{x \mid A\,x = b,\ 0 \leqq x\}$$

und

$$\mathfrak{X}_4 = \{x \mid A\,x = b,\ 0 \leqq x \leqq c\}$$

nur ganzzahlige Ecken.

Wir beschäftigen uns jetzt mit dem linearen Programm $\mathfrak{L}$.

$$\mathfrak{L}\text{: Suche } \max\{c'\,x \mid x \in \mathfrak{X}\}$$

wobei $\mathfrak{X} = \{x \mid A\,x = b,\ x \geqq 0\}$ ist:

LEMMA 2.4. *Wenn das Programm $\mathfrak{L}$ eine optimale Lösung besitzt, so besitzt es auch eine optimale Ecke.*

BEWEIS. Sei $\max\{c'\,x \mid x \in \mathfrak{X}\} = \lambda$ und sei

$$\mathfrak{X}_0 = \{x \mid A\,x = b,\ c'\,x = \lambda,\ x \geqq 0\}.$$

$\mathfrak{X}_0$ ist die Menge der optimalen Lösungen. Sie ist nach Voraussetzung nicht leer, besitzt also eine Ecke (einen Beweis, daß jede nicht leere Menge der Form $B\,x = c,\ x \geqq 0$ eine Ecke hat, findet man in [2, S. 84]). Wir zeigen, daß jede Ecke von $\mathfrak{X}_0$ auch eine Ecke von $\mathfrak{X}$ sein muß. Sei x eine Ecke von $\mathfrak{X}_0$, sei $x = (y+z)/2$ und seien y und z Punkte aus $\mathfrak{X}$. Dann gilt

$$c'\,y \leqq c'\,x = (c'\,y + c'\,z)/2 \quad \text{oder} \quad c'\,y \leqq c'\,z$$

und

$$c'\,z \leqq c'\,x = (c'\,y + c'\,z)/2 \quad \text{oder} \quad c'\,z \leqq c'\,y.$$

Es ist also $c'\,y = c'\,z = c'\,x$; y und z liegen in $\mathfrak{X}_0$, und da x eine Ecke von $\mathfrak{X}_0$ ist, so folgt $x = y = z$. Daher ist x auch eine Ecke von $\mathfrak{X}$.

Wir wenden jetzt das Vorhergehende auf das lineare Programm $\mathfrak{P}$ an.

$$\mathfrak{P}\text{: Suche } \max(c'\,x + d'\,y)$$

unter den Nebenbedingungen

$$\begin{aligned} A_1\,x + A_2\,y &\leqq a \\ B_1\,x + B_2\,y &= b \\ x &\geqq 0. \end{aligned}$$

Es gilt:

LEMMA 2.5. *Wenn a und b ganzzahlig sind, wenn $\begin{pmatrix} A_1, & A_2 \\ B_1, & B_2 \end{pmatrix}$ unimodular ist und wenn $\mathfrak{P}$ eine optimale Lösung besitzt, so besitzt $\mathfrak{P}$ auch eine ganzzahlige optimale Lösung.*

BEWEIS. Neben $\mathfrak{P}$ führen wir das lineare Programm $\mathfrak{R}$ ein.

$$\mathfrak{R}: \quad \text{Suche} \quad \max(c'x + d'y_1 - d'y_2)$$

nnter den Nebenbedingungen

$$\begin{pmatrix} A_1, & A_2, & -A_2, & I \\ B_1, & B_2, & -B_2, & 0 \end{pmatrix} \begin{pmatrix} x \\ y_1 \\ y_2 \\ z \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}, \quad \begin{pmatrix} x \\ y_1 \\ y_2 \\ z \end{pmatrix} \geqq 0.$$

Jede Lösung (zulässig oder optimal) von $\mathfrak{R}$ liefert eine Lösung von $\mathfrak{P}$, und aus jeder Lösung von $\mathfrak{P}$ kann man durch Aufspalten von y (in der Form $y = y_1 - y_2$ mit $y_1, y_2 \geqq 0$) Lösungen von $\mathfrak{R}$ gewinnen. Mit $\mathfrak{P}$ besitzt daher auch $\mathfrak{R}$ eine optimale Lösung. Wir beweisen das Lemma, indem wir zeigen, daß $\mathfrak{R}$ eine ganzzahlige optimale Lösung besitzt. Nach Lemma 2.2 ist mit

$$\begin{pmatrix} A_1, & A_2 \\ B_1, & B_2 \end{pmatrix} \quad \text{auch} \quad \begin{pmatrix} A_1, & A_2, & -A_2, & I \\ B_1, & B_2, & -B_2, & 0 \end{pmatrix}$$

unimodular, nach Satz 2.3 ist daher jede Ecke ganzzahlig, und nach Lemma 2.4 gibt es eine optimale Ecke. $\mathfrak{R}$ besitzt also eine ganzzahlige optimale Lösung. Daher besitzt auch $\mathfrak{P}$ eine ganzzahlige optimale Lösung.

Natürlich kann man eine dem Lemma 2.5 entsprechende Aussage auch über das zu $\mathfrak{P}$ duale Programm machen. Zusammen mit dem Dualitätssatz (Satz 1.1) ergibt sich dann:

SATZ 2.6. *Seien a, b, c und d ganzzahlige Vektoren und sei* $\begin{pmatrix} A_1, & A_2 \\ B_1, & B_2 \end{pmatrix}$ *eine unimodulare Matrix. Wir nehmen an, daß die durch die Systeme*

(i)
$$\begin{cases} A_1 x + A_2 y \leqq a \\ B_1 x + B_2 y = b \\ x \geqq 0 \end{cases}$$

und

(ii)
$$\begin{cases} A_1' u + B_1' v \geqq c \\ A_2' u + B_2' v = d \\ u \geqq 0 \end{cases}$$

definierten Mengen nicht leer sind. Dann gibt es auch ganzzahlige Punkte in den durch (i) *und* (ii) *definierten Mengen und es gilt:*

$$\max(c'x + d'y) = \min(a'u + b'v).$$

Dabei ist das Maximum (bzw. das Minimum) über die ganzzahligen Lösungen von (i) *(bzw. von* (ii)*) zu nehmen.*

3. Matrizen aus Nullen und Einsen. Wir bezeichnen mit $\mathfrak{M}$ die Menge der $m \times n$-Matrizen $(x_{i,k})$, welche den Bedingungen

$$\sum_{i=1}^{m} x_{i,k} \leqq r, \qquad k = 1, 2, \ldots, n,$$

$$\sum_{k=1}^{n} x_{i,k} \leqq s, \qquad i = 1, 2, \ldots, m,$$

$$0 \leqq x_{i,k} \leqq 1$$

genügen. Dabei seien r und s ganze positive Zahlen.

$\mathfrak{M}$ ist nicht leer, denn $x_{i,k} = 0$ (für alle i und k) liefert ein Element aus $\mathfrak{M}$.

Mit $\mathfrak{M}_0$ bezeichnen wir die Menge der Matrizen aus $\mathfrak{M}$, für welche $x_{i,k}$ nur die Werte 0 oder 1 annimmt.

SATZ 3.1. *$\mathfrak{M}$ ist die konvexe Hülle von $\mathfrak{M}_0$.*

BEWEIS. Die konvexe Menge $\mathfrak{M}$ hat nach Satz 2.1 nur ganzzahlige Ecken, diese können wegen $0 \leqq x_{i,k} \leqq 1$ nur aus Nullen und Einsen bestehen; die Ecken von $\mathfrak{M}$ liegen also in $\mathfrak{M}_0$. Jeder Punkt aus $\mathfrak{M}_0$ ist eine Ecke von $\mathfrak{M}$, denn wenn

$$(x_{i,k}) = \big((y_{i,k}) + (z_{i,k})\big)/2$$

ist, wobei $(y_{i,k})$ und $(z_{i,k})$ aus $\mathfrak{M}$ sind, so folgt aus $x_{i,k} = 0$ (aus $x_{i,k} = 1$), daß auch $y_{i,k} = z_{i,k} = 0$ (daß auch $y_{i,k} = z_{i,k} = 1$) ist. $\mathfrak{M}_0$ ist daher genau die Menge der Ecken von $\mathfrak{M}$, und da $\mathfrak{M}$ beschränkt ist, so wird es von seinen Ecken aufgespannt (s. z. B. [3]).

Wir definieren die Menge $\mathfrak{N}$ der Matrizen $(x_{i,k})$ durch

$$\sum_{i=1}^{m} x_{i,k} = r, \qquad k = 1, 2, \ldots, n,$$

$$\sum_{k=1}^{n} x_{i,k} = s, \qquad i = 1, 2, \ldots, m,$$

$$0 \leqq x_{i,k} \leqq 1.$$

Dabei seien r und s positive ganze Zahlen, und es gelte die Gleichung $sm = rn$.

$\mathfrak{N}$ ist nicht leer, denn $x_{i,k} = s/n = r/m$ (für alle i und k) liefert ein Element aus $\mathfrak{N}$.

Mit $\mathfrak{N}_0$ bezeichnen wir die Menge der Matrizen aus $\mathfrak{N}$, für welche $x_{i,k}$ nur die Werte 0 oder 1 annimmt. Genauso wie den Satz 3.1 beweist man den folgenden Satz.

SATZ 3.2. *$\mathfrak{N}$ ist die konvexe Hülle von $\mathfrak{N}_0$.*

Setzt man insbesondere $r = s = 1$ und $n = m$, so sind die Matrizen aus $\mathfrak{N}$ gerade die *doppelt stochastischen Matrizen*, und die Matrizen aus $\mathfrak{N}_0$ sind die *Permutationsmatrizen*. Der Satz 3.2 lautet dann:

KOROLLAR 3.3. *Die doppelt stochastischen Matrizen bilden die konvexe Hülle der Permutationsmatrizen.*

Korollar 3.3 wurde zuerst 1941 von G. BIRKHOFF in [1] bewiesen.

Sei $(d_{i,k})$ eine $m \times n$-Matrix, welche nur aus Nullen und Einsen besteht. Wir interessieren uns für das lineare Programm $\mathfrak{M}$.

$$\mathfrak{M}: \quad \text{Suche} \quad \lambda = \max \left\{ \sum_{i=1}^{m} \sum_{k=1}^{n} d_{i,k}\, x_{i,k} \,\middle|\, (x_{i,k}) \in \mathcal{M} \right\}.$$

Das dazu duale Programm $\mathfrak{M}'$ lautet

$$\mathfrak{M}': \quad \text{Suche} \quad \mu = \min \left\{ s \sum_{i=1}^{m} u_i + r \sum_{k=1}^{n} v_k + \sum_{i=1}^{m} \sum_{k=1}^{n} w_{i,k} \right\}$$

unter den Nebenbedingungen

$$(3.1) \quad u_i + v_k + w_{i,k} \geqq d_{i,k}, \quad u_i, v_k, w_{i,k} \geqq 0, \quad i = 1, 2, \ldots, m, \quad k = 1, 2, \ldots, n.$$

Offenbar gibt es zulässige Lösungen für $\mathfrak{M}'$, man setze z.B. alle u_i, v_k und $w_{i,k}$ gleich Eins.

Da $\mathfrak{M}$ und $\mathfrak{M}'$ zulässige Lösungen haben, so besitzen sie auch optimale Lösungen, und es gilt $\lambda = \mu$.

Die zur Definition der Menge $\mathcal{M}$ verwandte Matrix ist unimodular, daher ist der Satz 2.6 auf die dualen Programme $\mathfrak{M}$ und $\mathfrak{M}'$ anwendbar. Natürlich braucht man bei der Suche nach einem Minimum in $\mathfrak{M}'$ für u_i, v_k und $w_{i,k}$ nur die Werte 0 oder 1 in Betracht zu ziehen; größere ganze Zahlen machen die zu minimierende Linearform bestimmt unnötig groß. Deshalb läßt sich der Dualitätssatz für dieses Problem wie folgt formulieren:

SATZ 3.4. *Es ist*

$$\max \left\{ \sum_{i=1}^{m} \sum_{k=1}^{n} d_{i,k}\, x_{i,k} \right\} = \min \left\{ s \sum_{i=1}^{m} u_i + r \sum_{k=1}^{n} v_k + \sum_{i=1}^{m} \sum_{k=1}^{n} w_{i,k} \right\},$$

wobei $u_i, v_k, w_{i,k}$ und $x_{i,k}$ entweder 0 oder 1 sind und den Bedingungen $u_i + v_k + w_{i,k} \geqq d_{i,k}$ und $(x_{i,k}) \in \mathcal{M}$ unterworfen sind.

Der folgende Satz ist nur eine andere Formulierung dieser Aussage.

SATZ 3.5. *Sei $(d_{i,k})$ eine $m \times n$-Matrix aus Nullen und Einsen und seien r und s zwei ganze positive Zahlen. Die Maximalzahl der Einsen, welche in $(d_{i,k})$ so gefunden werden können, daß höchstens s in jeder Zeile und höchstens r in jeder Spalte stehen, ist gleich dem Minimum des Ausdrucks $sp + rq + a$, wobei p, q und a die folgende Bedingung erfüllen: es ist möglich, mit p Zeilen, q Spalten und a einzelnen Zellen alle Einsen von $(d_{i,k})$ zu überdecken.*

BEWEIS. Sei $(x_{i,k}) \in \mathcal{M}_0$. In $(x_{i,k})$ ist die Anzahl der Einsen je Zeile höchstens gleich s und die Anzahl der Einsen je Spalte höchstens gleich r. $\sum_{i=1}^{m} \sum_{k=1}^{n} d_{i,k}\, x_{i,k}$ zählt also Einsen aus $(d_{i,k})$, wobei höchstens s Einsen je Zeile und höchstens r Einsen je Spalte gezählt werden. Das Maximum dieser Anzahl ist λ. Nach Satz 3.4 ist $\lambda = \mu = \min(sp + rq + a)$, wobei

$$(3.2) \qquad p = \sum_{i=1}^{m} u_i, \qquad q = \sum_{k=1}^{n} v_k, \qquad a = \sum_{i=1}^{m} \sum_{k=1}^{n} w_{i,k}$$

gesetzt wurde. Wegen $u_i+v_k+w_{i,k}\geqq d_{i,k}$ werden die Einsen in $(d_{i,k})$ durch die p Zeilen (für deren Index die Gleichung $u_i=1$ gilt), durch die q Spalten (mit $v_k=1$) und durch die a Zeilen (mit $w_{i,k}=1$) überdeckt. Damit ist der Satz bewiesen.

Ist $r=1$, so kann man in der Definition von $\mathfrak{M}$ die Bedingung $x_{i,k}\leqq 1$ weglassen; sie ist sowieso erfüllt. Im dualen Programm $\mathfrak{M}'$ darf man dann $w_{i,k}=0$ (für alle Paare i, k) setzen. Aus Satz 3.4 wird damit

SATZ 3.6. *Es ist*

$$\max\left\{\sum_{i=1}^{m}\sum_{k=1}^{n} d_{i,k}\,x_{i,k}\right\}=\min\left\{s\sum_{i=1}^{m}u_i+\sum_{k=1}^{n}v_k\right\},$$

wobei u_i und v_k entweder 0 oder 1 sind und den Bedingungen $u_i+v_k\geqq d_{i,k}$ unterworfen sind. Das Maximum ist über alle $(x_{i,k})$ aus $\mathfrak{M}_0$ (mit $r=1$) zu nehmen.

Daraus folgt:

KOROLLAR 3.7. *Die Maximalzahl der Einsen in der Matrix $(d_{i,k})$, wobei höchstens s Einsen in jeder Zeile und höchstens eine Eins in jeder Spalte stehen, ist gleich dem Minimum des Ausdrucks $sp+q$, wobei mit p Zeilen und q Spalten alle Einsen aus $(d_{i,k})$ überdeckt werden können.*

Ist $r=s=1$, so gilt:

KOROLLAR 3.8. *Die Maximalzahl der Einsen in der Matrix $(d_{i,k})$ mit nicht mehr als einer Eins je Zeile oder Spalte ist gleich der Minimalzahl der Reihen (Zeilen oder Spalten), welche zusammen alle Einsen überdecken.*

Den Satz 3.6 mit $r=s=1$ aber beliebigen ganzzahligen $d_{i,k}\geqq 0$ findet man in [6, S. 67]. Das Korollar 3.8 wurde 1931 von D. KÖNIG bewiesen (s. [12, S. 240, Satz D]).

Für den nächsten Satz brauchen wir die Aussage

$\mathfrak{A}$: { Für die Anzahl a der im Durchschnitt von p beliebigen Zeilen und q beliebigen Spalten der Matrix $(d_{i,k})$ enthaltenen Einsen gilt $a\geqq sp+rq-rn$.

SATZ 3.9. *$\mathfrak{A}$ ist notwendig und hinreichend dafür, daß*

$$\lambda=\max\left\{\sum_{i=1}^{m}\sum_{k=1}^{n} d_{i,k}\,x_{i,k}\;\middle|\;(x_{i,k})\in\mathfrak{M}\right\}=sm$$

ist.

BEWEIS. 1. Es gelte $\mathfrak{A}$. Wir suchen $\mu=\min(sp+rq+a)$, wobei p, q und a wie in (3.2) definiert sind. u_i, v_k und $w_{i,k}$ sind entweder 0 oder 1 und genügen den Ungleichungen (3.1). Diese Ungleichungen sind für die p Indizes i mit $u_i=1$ und für die q Indizes k mit $v_k=1$ erfüllt, gleich welchen Wert $w_{i,k}$ hat. Für die restlichen $(m-p)(n-q)$ Indexpaare (i, k), welche dabei noch nicht berücksichtigt wurden, gilt: wenn $d_{i,k}=1$ ist, so muß auch $w_{i,k}=1$ sein. Daher ist

$$a=\sum_{i=1}^{m}\sum_{k=1}^{n} w_{i,k}$$

größer oder gleich der Anzahl der $d_{i,k}$, welche im Durchschnitt der $(m-p)$ Zeilen und $(n-q)$ Spalten liegen und gleich Eins sind. Wegen $\mathfrak{A}$ gilt also

$$a \geqq s(m-p)+r(n-q)-rn$$

oder

$$sp+rq+a \geqq sm.$$

Damit ist gezeigt, daß $\mu=\min(sp+rq+a) \geqq sm$ ist. Setzt man aber alle $u_i=1$ und alle v_k und $w_{i,k}$ gleich 0, so ist (3.1) erfüllt, und die Linearform $sp+rq-a$ wird gleich sm. Also ist $\mu=sm$ und nach dem Dualitätssatz auch $\lambda=sm$.

2. Sei nun $\lambda=sm=\sum_{i=1}^{m}\sum_{k=1}^{n} d_{i,k}x_{i,k}$ und $(x_{i,k})\in\mathfrak{M}_0$. Wegen $\sum_{k=1}^{n} x_{i,k}\leqq s$ und $\sum_{i=1}^{m}\sum_{k=1}^{n} x_{i,k}\leqq sm$ muß dann $\sum_{i=1}^{m}\sum_{k=1}^{n} x_{i,k}=sm$ und $\sum_{k=1}^{n} x_{i,k}=s$ sein. Ferner folgt: wenn $x_{i,k}=1$ ist, so ist auch $d_{i,k}=1$. Es reicht also, die Bedingung $\mathfrak{A}$ für die Matrix $(x_{i,k})$ zu beweisen, sie gilt dann auch für die Matrix $(d_{i,k})$.

Für jede beliebige Matrix aus Nullen und Einsen gilt: Anzahl aller „1" in der Matrix + Anzahl der „1" im Durchschnitt von p Zeilen und q Spalten ist größer oder gleich der Anzahl der „1" in den p Zeilen + Anzahl der „1" in den q Spalten.

In unserem Falle erhält man daraus (angewandt auf $(x_{i,k})$)

$$sm+a \geqq sp+x. \tag{3.3}$$

Hierbei ist

sm = Anzahl aller „1" in $(x_{i,k})$,
a = Anzahl der „1" im Durchschnitt von p Zeilen und q Spalten,
sp = Anzahl der „1" in den p Zeilen,
x = Anzahl der „1" in den q Spalten.

Sei nun y die Anzahl der „1" in den komplementären $(n-q)$ Spalten. Dann gilt

$$x+y=sm \quad \text{und} \quad y \leqq r(n-q),$$

denn die Summe der „1" pro Spalte ist höchstens gleich r. Aus diesen Bedingungen und (3.3) folgt die Ungleichung $a \geqq sp+rq-rn$. Damit ist $\mathfrak{A}$ hergeleitet.

Im Sonderfall $r=1$ braucht man die Bedingung

$\mathfrak{B}$: *Bei irgend p Zeilen der Matrix $(d_{i,k})$ stehen in wenigstens sp verschiedenen Spalten Einsen.*

Satz 3.10. *$\mathfrak{B}$ ist notwendig und hinreichend dafür, daß*

$$\lambda=\max\left\{\sum_{i=1}^{m}\sum_{k=1}^{n} d_{i,k}x_{i,k} \;\middle|\; (x_{i,k})\in\mathfrak{M}\right\}=sm$$

ist. Hierbei ist $\mathfrak{M}$ mit $r=1$ definiert.

BEWEIS. 1. Es gelte $\mathfrak{B}$. Wir wenden den Satz 3.6 an. Sei $\mu = \min(sp+q)$, wobei p und q wie in (3.2) definiert sind. Die u_i und v_k sind entweder 0 oder 1 und genügen den Ungleichungen $u_i + v_k \geqq d_{i,k}$. Da für die p Indizes i mit $u_i = 1$ diese Ungleichungen bestimmt erfüllt sind, untersuchen wir die $(m-p)$ komplementären Indizes. In den entsprechenden Zeilen gibt es (wegen $\mathfrak{B}$) wenigstens $s(m-p)$ Spalten, welche Einsen enthalten. Daher muß $q \geqq s(m-p)$ sein. Es folgt: $sp+q \geqq sm$, und genau wie beim Beweise von Satz 3.9 schließen wir, daß $sm = \min(sp+q) = \lambda$ ist.

2. Sei nun $\lambda = sm$. Für eine optimale Matrix $(x_{i,k})$ aus $\mathfrak{M}_0$ gilt dann (vgl. den Beweis von Satz 3.9)

$$\sum_{i=1}^{m} \sum_{k=1}^{n} x_{i,k} = sm \quad \text{und} \quad \sum_{k=1}^{n} x_{i,k} = s.$$

Für irgend p Zeilen (Menge der dazu gehörenden Indizes $= \mathfrak{I}$) ist also

$$\sum_{i \in \mathfrak{I}} \sum_{k=1}^{n} x_{i,k} = sp.$$

In jeder Spalte der Matrix $(x_{i,k})$ steht höchstens eine Eins, also ist sp gleich der Anzahl der Spalten, für welche in einer der p Zeilen eine Eins steht. Da $(d_{i,k})$ höchstens zusätzliche Einsen gegenüber $(x_{i,k})$ enthält, so folgt: sp ist kleiner oder gleich der Anzahl der Spalten der Matrix $(d_{i,k})$, welche in einer der p Zeilen eine Eins haben. Damit ist Satz 3.10 bewiesen.

Sei jetzt $r = s = 1$.

KOROLLAR 3.11. *Notwendig und hinreichend dafür, daß man in der Matrix $(d_{i,k})$ m Einsen finden kann, derart, daß höchstens eine Eins je Zeile oder Spalte auftritt, ist die Bedingung: Irgend p verschiedene Zeilen haben in wenigstens p verschiedenen Spalten eine Eins stehen.*

4. Vertreter-Systeme. Wir wollen nun die Sätze des vorhergehenden Abschnittes anwenden. Zuerst erklären wir, was wir unter einem Vertreter-System verstehen wollen.

Sei $\mathfrak{A} = \{a_1, a_2, \ldots, a_n\}$ eine endliche Menge mit den n Elementen a_k. $\mathfrak{A}_i$ und $\mathfrak{S}_i$ $(i = 1, 2, \ldots, m)$ seien Teilmengen von $\mathfrak{A}$. Wir nennen die Mengen $\mathfrak{S}_i$ ein (r, s)-Vertreter-System der Mengen $\mathfrak{A}_i$, wenn folgende drei Bedingungen erfüllt sind.

(i) $\mathfrak{S}_i \subseteq \mathfrak{A}_i$ (die $\mathfrak{S}_i$ sind „Vertreter" der Mengen $\mathfrak{A}_i$).

(ii) $\mathfrak{S}_i$ enthält genau s Elemente.

(iii) Jedes Element a_k kommt in höchstens r verschiedenen Mengen $\mathfrak{S}_i$ vor.

Aus den Bedingungen folgt: Bei einem $(1, s)$-Vertreter-System sind die Mengen $\mathfrak{S}_i$ paarweise elementefremd.

Für ein $(1, 1)$-Vertreter-System lassen sich die Bedingungen (i), (ii) und (iii) zusammenfassen zu

(iv) Zu jeder Menge $\mathfrak{A}_i$ gibt es ein Element als Vertreter; die Vertreter verschiedener Mengen $\mathfrak{A}_i$ sind voneinander verschieden.

(1, 1)-Vertreter-Systeme wurden zuerst 1935 von PH. HALL in [7] betrachtet.

Wir definieren nun eine Matrix $(d_{i,k})$, welche wir dann in die Sätze 3.9 und 3.10 einsetzen wollen.

$$d_{i,k}=1 \quad \text{für} \quad a_k \in \mathscr{A}_i,$$
$$d_{i,k}=0 \quad \text{für} \quad a_k \notin \mathscr{A}_i.$$

SATZ 4.1. *Wenn für ein* $(x_{i,k}) \in \mathscr{M}_0$ *die Beziehung*

$$s\,m=\sum_{i=1}^{m}\sum_{k=1}^{n} d_{i,k}\,x_{i,k} \tag{4.1}$$

besteht, so gibt es ein (r, s)*-Vertreter-System. Existiert ein* (r, s)*-Vertreter-System, so gibt es eine Matrix* $(x_{i,k}) \in \mathscr{M}_0$*, für die* (4.1) *gilt.*

BEWEIS. 1. Sei $(x_{i,k}) \in \mathscr{M}_0$ gegeben und gelte (4.1). Wir definieren $\mathscr{S}_i$ als die Menge derjenigen a_k, für deren Index k die Gleichung $x_{i,k}=1$ gilt. Aus (4.1) folgt: Wenn $x_{i,k}=1$ ist, so ist auch $d_{i,k}=1$ (s. den Beweis von Satz 3.9), also folgt aus $x_{i,k}=1$, daß $a_k \in \mathscr{A}_i$ liegt. Dies liefert: (i) $\mathscr{S}_i \subseteq \mathscr{A}_i$.

Da $\sum_{k=1}^{n} x_{i,k}=s$ ist, so enthält $\mathscr{S}_i$ genau s Elemente. Damit ist (ii) gezeigt.

Aus $\sum_{i=1}^{m} x_{i,k} \leqq r$ folgt, daß a_k in höchstens r verschiedenen $\mathscr{S}_i$ liegen kann; also gilt auch (iii).

2. Sei nun ein (r, s)-Vertreter-System gegeben. Wir setzen $x_{i,k}=1$, wenn a_k in $\mathscr{S}_i$ liegt. Sonst sei $x_{i,k}=0$. Wegen (ii) hat man $\sum_{k=1}^{n} x_{i,k}=s$ und wegen (iii) hat man $\sum_{i=1}^{m} x_{i,k} \leqq r$. Da wegen (i) aus $x_{i,k}=1$ die Gleichung $d_{i,k}=1$ folgt, so ist

$$\sum_{k=1}^{n} d_{i,k}\,x_{i,k}=s \quad \text{und} \quad \sum_{i=1}^{m}\sum_{k=1}^{n} d_{i,k}\,x_{i,k}=s\,m.$$

Damit ist Satz 4.1 bewiesen.

In Verbindung mit den Überlegungen des vorigen Abschnitts ergeben sich nun die folgenden Sätze.

SATZ 4.2. *Sei* $\mathscr{T}$ *eine Teilmenge von* $\mathscr{A}$ *und* $\mathscr{I}$ *eine Teilmenge von* $\{1, 2, \ldots, m\}$. *Mit* $|\mathscr{T}|$ (*bzw.* $|\mathscr{I}|$) *bezeichnen wir die Anzahl der Elemente in* $\mathscr{T}$ (*bzw. in* $\mathscr{I}$). *Mit dieser Bezeichnungsweise gilt: Notwendig und hinreichend für die Existenz eines* (r, s)*-Vertreter-Systems ist, daß für jedes* $\mathscr{T}$ *und für jedes* $\mathscr{I}$

$$\sum_{i \in \mathscr{I}} |\mathscr{A}_i \cap \mathscr{T}| \geqq s|\mathscr{I}| + r|\mathscr{T}| - r\,n$$

ist.

SATZ 4.3. *Notwendig und hinreichend für die Existenz eines* $(1, s)$*-Vertreter-Systems ist die Bedingung: Irgend* p *verschiedene Mengen* $\mathscr{A}_i$ *enthalten zusammen wenigstens* sp *verschiedene Elemente.*

KOROLLAR 4.4. *Notwendig und hinreichend für die Existenz eines* $(1, 1)$*-Vertreter-Systems ist, daß irgend* p *verschiedene Mengen* $\mathscr{A}_i$ *unter sich wenigstens* p *verschiedene Elemente enthalten.*

Mit Rücksicht auf den Satz 4.1 folgt der Satz 4.2 unmittelbar aus dem Satz 3.9, der Satz 4.3 folgt unmittelbar aus dem Satz 3.10, und das Korollar 4.4' folgt aus dem Korollar 3.11.

Korollar 4.4 ist der klassische Satz über Vertreter-Systeme. Er gilt auch für unendliche Mengen. Man sehe dazu [7] und [6].

Benutzt man statt der Matrix $(d_{i,k})$ die dazu transponierte Matrix, so erhält man sofort

SATZ 4.5. *Notwendig und hinreichend für die Existenz eines $(r, 1)$-Vertreter-Systems ist die Bedingung: Irgend p verschiedene Elemente haben mit wenigstens rp verschiedenen Mengen $\mathfrak{A}_i$ einen nicht leeren Durchschnitt.*

Wir geben noch eine Verallgemeinerung der Sätze 3.9 und 4.2 an, den Beweis führen wir nicht aus, da er vollkommen analog zum Vorhergehenden verläuft.

Seien s_i $(i=1, 2, \ldots, m)$ und r_k $(k=1, 2, \ldots, n)$ ganze positive Zahlen und sei die konvexe Menge $\mathfrak{M}$ der $m \times n$-Matrizen $(x_{i,k})$ definiert durch

$$\sum_{i=1}^{m} x_{i,k} \leqq r_k$$

$$\sum_{k=1}^{n} x_{i,k} \leqq s_i$$

$$0 \leqq x_{i,k} \leqq 1.$$

SATZ 4.6. *Für die Gültigkeit der Gleichung*

$$\max\left\{\sum_{i=1}^{m}\sum_{k=1}^{n} d_{i,k}\, x_{i,k} \,\middle|\, (x_{i,k}) \in \mathfrak{M}\right\} = \sum_{i=1}^{m} s_i$$

ist die folgende Bedingung notwendig und hinreichend. Die aus Nullen und Einsen bestehende Matrix $(d_{i,k})$ hat die Eigenschaft, daß die Anzahl der im Durchschnitt der Zeilen Nr. $i_1, i_2, \ldots, i_p$ und der Spalten Nr. $k_1, k_2, \ldots, k_q$ stehenden Einsen größer oder gleich ist dem Ausdruck

$$\sum_{\nu=1}^{p} s_{i_\nu} + \sum_{\nu=1}^{q} r_{k_\nu} - \sum_{k=1}^{n} r_k.$$

DEFINITION. Ein allgemeines Vertreter-System für die Teilmengen $\mathfrak{A}_i$ $(i=1, 2, \ldots, m)$ der endlichen Menge $\mathfrak{A}$ ist ein System von Teilmengen $\mathfrak{S}_i$ mit folgenden Eigenschaften: $\mathfrak{S}_i \subseteq \mathfrak{A}_i$, $|\mathfrak{S}_i| = s_i$ und das Element a_k kommt in höchstens r_k der Mengen $\mathfrak{A}_i$ vor.

Ein allgemeines Vertretersystem wird also mit Hilfe der ganzen positiven Zahlen s_i $(i=1, 2, \ldots, m)$ und r_k $(k=1, 2, \ldots, n)$ definiert.

SATZ 4.7. *Notwendig und hinreichend für die Existenz eines allgemeinen Vertreter-Systems (welches zu den Zahlen s_i und r_k gehört) ist die folgende Bedingung:*

Wenn $\mathfrak{R} = \{a_{k_1}, a_{k_2}, \ldots, a_{k_q}\}$ eine beliebige Teilmenge von $\mathfrak{A}$ ist und wenn $\mathfrak{J}$ eine beliebige Teilmenge von $\{1, 2, \ldots, m\}$ ist, so gilt

$$\sum_{i \in \mathfrak{J}} |\mathfrak{A}_i \cap \mathfrak{R}| \geqq \sum_{i \in \mathfrak{J}} s_i + \sum_{\nu=1}^{q} r_{k_\nu} - \sum_{k=1}^{n} r_k.$$

Literatur

[1] Birkhoff, G.: Tres observationes sobre el algebra lineal. Rev. Univ. Nac. Tucuman, Ser. A **6**, 147 (1941).
[2] Burger, E.: Einführung in die Theorie der Spiele. Berlin 1959.
[3] Charnes, A., and W. W. Cooper: The Minkowski-Weyl-Farkas Theorem. Proc. nat. Acad. Sci., Wash. **44**, 914–916 (1958).
[4] Ford jr., L. R., and D. R. Fulkerson: Network Flow and Systems of Representatives. Canad. J. Math. **10**, 78–84 (1958).
[5] Goldman, A. J., and A. W. Tucker: Theory of Linear Programming, in [13], 53–97.
[6] Hall jr., Marshall: Theorems on Choice. In: Some Aspects of Analysis and Probability. (Surveys in Applied Mathematics IV.) New York 1958.
[7] Hall, Ph.: On Representatives of Subsets. J. London Math. Soc. **10**, 26–30 (1935).
[8] Heller, I., and C. B. Tompkins: An Extension of a Theorem of Dantzig's, in [13], 247–254.
[9] Hoffman, A. J., and J. B. Kruskal: Integral Boundary Points of Convex Polyhedrons, in [13], 223–246.
[10] Hoffman, A. J., and H. W. Kuhn: On Systems of Representatives, in [13], 199–206.
[11] Hoffman, A. J., and H. W. Kuhn: Systems of Distinct Representatives and Linear Programming. Amer. Math. Monthly **63**, 455–460 (1956).
[12] König, D.: Theorie der endlichen und unendlichen Graphen. Leipzig 1936.
[13] Kuhn, H. W., and A. W. Tucker (Ed.): Linear Inequalities and Related Systems. Ann. Math. Studies **38** (1956).

Tübingen, Mathematisches Institut der Universität

(Eingegangen am 1. Oktober 1960)

SCHUBERT, H.
Math. Zeitschr. 76, 116–148 (1961)

Bestimmung der Primfaktorzerlegung von Verkettungen

Von
HORST SCHUBERT

In Verallgemeinerung des Sachverhaltes für Knoten wurde von Y. HASHIZUME [2] gezeigt, daß jede nicht-zerfallende Verkettung in der 3-Sphäre S^3 eine eindeutige Zerlegung in Primfaktoren besitzt (vgl. 6.1 im folgenden). Nachdem von W. HAKEN [1] endliche Verfahren angegeben wurden, von einem Knoten zu entscheiden, ob er ein Kreis ist, und von einer Verkettung zu entscheiden, ob sie zerfällt, läßt sich ein endliches Verfahren angeben, das die Primfaktorzerlegung einer Verkettung herstellt.

Es wird hier zunächst das Verfahren von HAKEN beschrieben. Durch eine Variante des Reduktionsansatzes (3.2) ergibt sich eine erhebliche Vereinfachung und außerdem die Anwendbarkeit des Verfahrens auf Flächen in nicht-orientierbaren Mannigfaltigkeiten. Zur Illustration ist in § 4 die einfachste Anwendung ausgeführt, nämlich die Entscheidung, ob eine Verkettung in der S^3 zerfällt. Danach werden in § 5 Resultate von HAKEN verallgemeinert zu einem Bestimmungsverfahren für das Geschlecht von Knoten in beliebigen 3-dimensionalen, kompakten Mannigfaltigkeiten. Schließlich wird in § 6 die Faktorzerlegung von Verkettungen in der S^3 behandelt.

Die Endlichkeit der Verfahren bedeutet nicht, daß sie praktisch ausführbar wären. Man kann aber zeigen, daß sie finit im strengen Sinne sind, d.h. daß sich die Anzahl der benötigten Schritte nach oben abschätzen läßt. Wir gehen hierauf nicht ein, weil die uns vorliegenden Abschätzungen so grob sind, daß sie keine Information über die Anzahl der tatsächlich benötigten Schritte ergeben. Außerdem haben wir darauf verzichtet, das Normalgleichungssystem in 2.4 möglichst klein zu halten, was durch Änderungen der Zellenzerlegung und durch Verschärfung des 8. Schrittes in 2.2 möglich wäre. Dies scheint unwesentlich zu sein, solange das Verfahren nicht tatsächlich ausgeführt werden soll. Wir haben uns daher für die einfachere Beschreibung entschieden.

§ 1. Vorbemerkungen zum Verfahren von HAKEN

1.1. Bezeichnungen. Der 3-dimensionale euklidische Raum wird stets mit R^3, die 3-Sphäre mit S^3 bezeichnet. M sei eine dreidimensionale kompakte Mannigfaltigkeit, die möglicherweise berandet ist. $\dot{M}$ bezeichne den Rand von M. M sei mit einer festen regulären Zellenzerlegung Σ versehen. Für die Zellen von Σ sei eine affine Struktur fixiert, so daß es Sinn hat, von geradlinigen Unterteilungen zu sprechen. Alle vorkommenden Flächen sollen sim-

plizial in einer jeweils geeigneten, geradlinigen Unterteilung von Σ sein. Unter Flächen sind (abweichend von HAKEN) stets zusammenhängende Flächen zu verstehen, andernfalls sprechen wir von Flächensystemen. Der Rand einer Fläche kann unzusammenhängend sein. $\mathfrak{F}$ sei stets eine kompakte Fläche in M, deren Rand leer ist oder auf $\dot{M}$ liegt. $\dot{\mathfrak{F}}$ bezeichne den Rand von $\mathfrak{F}$. Die Charakteristik[1]) von $\mathfrak{F}$ wird mit $\chi(\mathfrak{F})$ bezeichnet. Unter isotopen Deformationen sind stets semilineare, isotope Deformationen zu verstehen.

1.2. Überblick. Es wird sich darum handeln, Flächen zu betrachten, die für M „charakteristisch" sind. Was darunter zu verstehen ist, hängt von der jeweils vorliegenden Anwendung des Verfahrens ab. Zur Fixierung der Ideen sei C eine einfache Kurve auf $\dot{M}$, die auf M nullhomolog modulo 2 ist. Außerdem sei C Unterkomplex von $\dot{M}$ (bezüglich Σ). Es gibt dann Flächen auf M, die von C berandet werden, und unter ihnen solche von maximaler Charakteristik[1]). Diese sind die charakteristischen Flächen für das Verfahren, solche Flächen zu bestimmen (vgl. 5.5). Im allgemeinen wird unter diesen Flächen keine sein, die Unterkomplex von M ist. Man muß sich daher mit der Forderung begnügen, daß eine solche Fläche bezüglich Σ in „schöner Lage" ist. „Schöne Lage" wird in 2.3 durch den Begriff der Normalfläche präzisiert, wobei die zu Σ gehörige Normalzerlegung Σ' von M (2.1) eingeht.

Die Normalflächen sind durch ihren Verlauf in den Umgebungen der Ecken von Σ vollständig beschrieben (bis auf unwesentliche Isotopie). Hierauf beruht es, daß sie gewissen nicht-negativen ganzen Lösungen eines homogenen linearen Gleichungssystems entsprechen (2.4). Die nicht-negativen ganzen Lösungen dieses „Normalgleichungssystems" sind nicht-negative ganze Linearkombinationen von endlich vielen Fundamentallösungen (2.5). Zu Fundamentallösungen gehörige Flächen werden als Fundamentalflächen bezeichnet.

Durch ein Normalisierungsverfahren (2.2), das über isotope Deformationen hinausgeht, werden charakteristische Flächen in Normalflächen verwandelt. Für jede Anwendung ist gesondert nachzuprüfen, daß dabei wieder charakteristische Flächen entstehen, wobei die Existenz charakteristischer Flächen hypothetisch sein kann[2]). Ferner wird nachgewiesen, daß charakteristische Fundamentalflächen vorhanden sind, wenn es charakteristische Normalflächen gibt. Dieser Nachweis hängt wieder von der jeweiligen Problemstellung ab, beruht aber auf einem allgemeinen Reduktionsansatz, der in § 3 dargestellt ist. Vermöge der Fundamentalflächen erhält man nun ein endliches Verfahren für die Bestimmung charakteristischer Flächen, wobei zugleich über die Existenz solcher Flächen entschieden wird. Dazu muß allerdings noch für jede Fundamentalfläche feststellbar sein, ob sie charakteristisch ist oder nicht, was trivial (wie in § 4 und § 5) oder Gegenstand eines anderen Verfahrens (wie in § 6) sein kann.

[1]) In Angleichung an den Gebrauch bei beliebigen Polyedern ist das Vorzeichen der Charakteristik so gewählt, daß die 2-Sphäre die Charakteristik + 2, also größte Charakteristik hat.

[2]) Dies ist beispielsweise der Fall, wenn man oben für die Kurve C nicht voraussetzt, daß sie auf M nullhomolog modulo 2 ist .

§ 2. Normalflächen

2.1. Normalzerlegung. Zur Zellenzerlegung Σ von M wird eine zugehörige Normalzerlegung Σ' hergestellt und zwar derart, daß jeder Zelle von Σ (der Dimension 0, 1, 2, 3) eine 3-Zelle von Σ' zugeordnet ist (Fig. 1). Zunächst werden die Ecken von Σ mit paarweise disjunkten *Kugel*umgebungen (P-Raumstücke)[3]) versehen, danach werden die Kantenstücke von Σ außerhalb der eben gewählten Kugeln mit paarweise disjunkten *Balken*umgebungen (K-Raumstücke)[3]) versehen. Zur Veranschaulichung merken wir an, daß das jetzt vorhandene Kugel- und Balkensystem auf ein Heegard-Diagramm von M führt. Nunmehr werden die Stücke der 2-Zellen von Σ außerhalb des

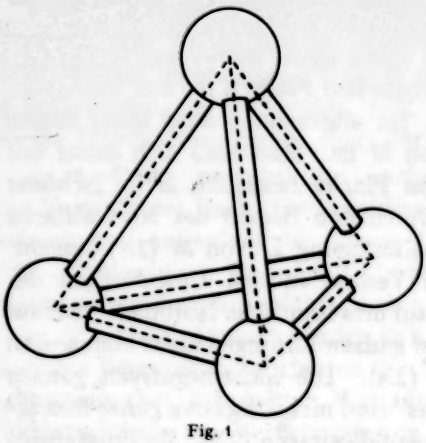

Fig. 1

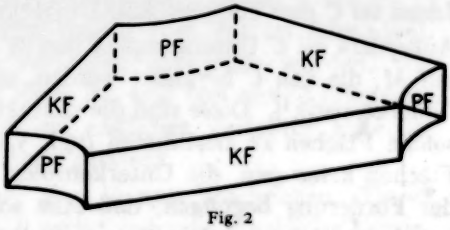

Fig. 2

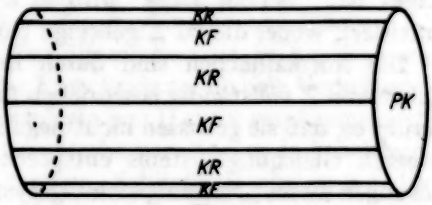

Fig. 3

Kugel- und Balkensystems mit paarweise disjunkten *Platten*umgebungen (F-Raumstücke)[3]) versehen. (Diese Platten sind in Fig. 1 nicht mitgezeichnet.) Es verbleiben noch *Restraumstücke* (R-Raumstücke)[3]) im Innern der 3-Zellen von Σ. Damit sind die 3-Zellen von Σ' beschrieben. Sie sollen natürlich „schöne" Umgebungen[4]) der entsprechenden Zellenteile von Σ sein, etwa simpliziale Umgebungen in genügend feinen simplizialen Unterteilungen von Σ.

Σ' ist durch Σ bis auf Isomorphie bestimmt; das Schema von Σ' kann aus dem Schema von Σ abgelesen werden.

Außer den 3-Zellen werden immer wieder die 2-Zellen von Σ' benötigt. Wir beschreiben daher ihre gegenseitige Lage auf den Rändern der 3-Zellen, wobei wir von den später nicht weiter benötigten Restraumstücken absehen.

Auf dem *Rande einer Platte* (Fig. 2) liegen Elementarflächenstücke, in denen Kugeln und Balken an die Platte anstoßen (PF und KF in Fig. 2). Diese Flächenstücke setzen sich zu einem Kreisring zusammen, auf dem Kugel- und Balkenanstöße abwechseln. In den beiden restlichen Elementarflächenstücken (Boden und Dach der Platte) stoßen entweder zwei Restraumstücke

[3]) Bezeichnung bei HAKEN [1].

[4]) Auf Einzelheiten dieser Art gehen wir hier und im folgenden nicht ein. Sie sind bei HAKEN ausführlich dargestellt.

an, oder es liegt eines von beiden auf $\dot{M}$, während am anderen ein Restraumstück anstößt.

Auf dem *Rande eines Balkens* (Fig. 3) liegen zwei Flächenstücke, in denen Kugeln anstoßen (PK in Fig. 3). Der dazwischen liegende Kreisring ist in Streifen zerlegt, in denen abwechselnd Platten und Restraumstücke anstoßen (KF und KR in Fig. 3). Höchstens ein Streifen kann auf $\dot{M}$ liegen; er liegt dann zwischen zwei Streifen, in denen Platten anstoßen (statt eines mit KR bezeichneten).

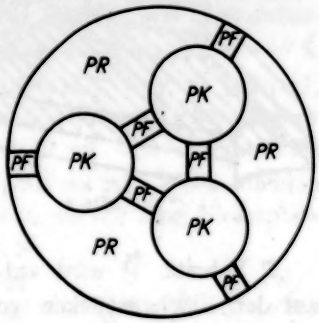

Fig. 4

Auf dem *Rande einer Kugel* (Fig. 4) liegen als „Inseln" Flächenstücke, in denen Balken anstoßen (PK in Fig. 4). Diese Inseln sind verbunden durch „Brücken" (PF), in denen Platten anstoßen. Es verbleiben „Seen" (PR), in denen Restraumstücke anstoßen. Es kann auch sein, daß ein See auf $\dot{M}$ liegt, wir sprechen dann von einem „Randsee".

Je nach Art der zusammenstoßenden 3-Zellen können die 2-Zellen von Σ' als PK-, PF-, PR-, KF-, KR-, FR-Flächenstücke bezeichnet werden (wie dies in den Figuren geschehen ist) bzw. als $P\dot{M}$-, $K\dot{M}$-, $F\dot{M}$-Flächenstücke, wenn sie auf $\dot{M}$ liegen. Beispielsweise sind Inseln PK-Flächenstücke, Randseen $P\dot{M}$-Flächenstücke.

2.2. Normalisieren charakteristischer Flächen. Es sei $\mathfrak{F}$ eine für M charakteristische Fläche. (Die Definition dieser Flächen hängt von der jeweiligen Anwendung ab.) Der Rand $\dot{\mathfrak{F}}$ von $\mathfrak{F}$ liege auf $\dot{M}$. Er kann leer sein. $\mathfrak{F}$ soll durch eine Fläche ersetzt werden, die bezüglich Σ' in übersichtlicher Lage ist. Dies wird durch isotope Deformation und zwei Zerschneidungsoperationen (7. und 8. Schritt) erreicht. Dabei nehmen wir an, daß die Definition der charakteristischen Flächen so beschaffen ist, daß isotope semilineare Deformationen erlaubt sind und daß keine Komponente von $\dot{\mathfrak{F}}$ ein Elementarflächenstück auf $\dot{M}$ berandet, was in den Anwendungen stets der Fall sein wird.

1. Schritt. Durch isotope Deformation von $\mathfrak{F}$ bei festem Rand $\dot{\mathfrak{F}}$ wird erreicht, daß $\dot{\mathfrak{F}} = \mathfrak{F} \cap \dot{M}$ ist. Der übrige Durchschnitt von $\mathfrak{F}$ und $\dot{M}$ wird von $\dot{M}$ abgehoben[5]).

2. Schritt. $\mathfrak{F}$ wird durch eine Isotopie bei festem Rand aus allen Restraumstücken ausgefegt.

3. Schritt. $\dot{\mathfrak{F}}$ wird auf $\dot{M}$ isotop aus denjenigen Flächenstücken ($F\dot{M}$) herausdeformiert, in denen Platten anstoßen. Diese Isotopie von $\dot{\mathfrak{F}}$ wird zu einer Isotopie von $\mathfrak{F}$ so fortgesetzt, daß der Durchschnitt von $\mathfrak{F}$ mit den Restraumstücken leer bleibt.

[5]) Die Einzelheiten hierfür und für die folgenden Schritte findet man in [5] und [6].

4. Schritt. $\mathfrak{F}$ wird durch eine Isotopie bei festem Rand in den Platten „straffgezogen", so daß $\mathfrak{F}$ in den Platten Elementarflächenstücke bildet, die „parallel" zu Boden und Dach liegen (Fig. 5, vgl. Fig. 2). Der Rand eines solchen Flächenstückes bildet je einen Querschnitt auf den PF- und KF-Flächenstücken im Rande der Platte. Das Straffziehen von $\mathfrak{F}$ in einer Platte $\mathfrak{P}$ erfolgt so, daß eine geeignete „senkrechte Säule" in $\mathfrak{P}$ zu $\mathfrak{P}$ aufgeblasen wird. $\mathfrak{F}$ wird also zum Teil aus $\mathfrak{P}$ in die anstoßenden Kugeln und Balken verdrängt. Der Durchschnitt von $\mathfrak{F}$ mit den Restraumstücken bleibt dabei leer.

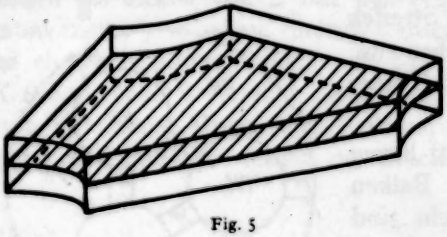

Fig. 5

5. Schritt. $\dot{\mathfrak{F}}$ wird auf $\dot{M}$ straffgezogen. Es kann sein (Fig. 6), daß $\dot{\mathfrak{F}}$ auf den Flächenstücken von $\dot{M}$, in denen Kugeln ($P\dot{M}$) und Balken ($K\dot{M}$) anstoßen, „Schleifen" bildet, indem ein Bogen von $\dot{\mathfrak{F}}$ auf einem $P\dot{M}$-Flächenstück ein anstoßendes $K\dot{M}$-Flächenstück mit sich selbst verbindet oder indem ein Bogen von $\dot{\mathfrak{F}}$ auf einem $K\dot{M}$-Flächenstück ein anstoßendes $P\dot{M}$-Flächenstück mit sich selbst verbindet. Solche Schleifen von $\dot{\mathfrak{F}}$ können der Reihe nach durch isotope Deformation von $\dot{\mathfrak{F}}$ beseitigt werden, so daß danach $\dot{\mathfrak{F}}$ auf den $P\dot{M}$- und $K\dot{M}$-Flächenstücken Querschnitte bildet, diejeweils verschiedene anstoßende Flächenstücke verbinden[6]). Die Deformation von $\dot{\mathfrak{F}}$ wird so zu einer isotopen Deformation von $\mathfrak{F}$ fortgesetzt, daß sich am Durchschnitt von $\mathfrak{F}$ mit den Restraumstücken und Platten nichts ändert.

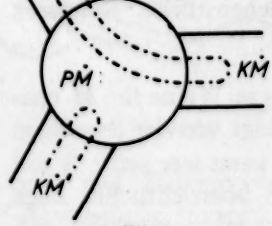

Fig. 6

6. Schritt. $\mathfrak{F}$ wird bei festem Rand in den Balken straffgezogen durch eine Isotopie, die am Verlauf von $\mathfrak{F}$ in den Platten nichts ändert. Dies geschieht so, daß Teile von $\mathfrak{F}$ aus den Balken in die anstoßenden Kugeln verdrängt werden, so daß $\mathfrak{F}$ danach in den Balken Bänder und vielleicht auch Zylinder bildet, die in Balkenrichtung verlaufen (Fig. 7, vgl. Fig. 3). Wegen des 5. Schrittes ist das Straffziehen auch in denjenigen Balken möglich, die an $\dot{M}$ anstoßen.

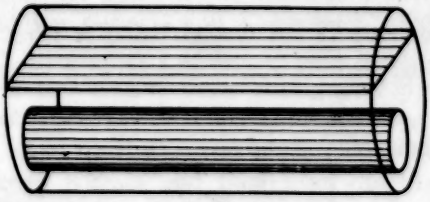

Fig. 7

7. Schritt. Vereinfachung in den Kugeln. Es sei $\mathfrak{K}$ eine Kugel von Σ' (P-Raumstück) und $\dot{\mathfrak{K}}$ ihr Rand. Falls $\mathfrak{K}$ an $\dot{M}$ anstößt, sind die Bögen von

[6]) Wir erinnern daran, daß keine Komponente von $\dot{\mathfrak{F}}$ ein Elementarflächenstück auf $\dot{M}$ berandet.

$\dot{\mathfrak{F}}$ auf $\dot{\mathfrak{K}}$ zu den Schnitten von $\dot{\mathfrak{K}}$ und $\mathfrak{F}$ zu rechnen. Unter dieser Vereinbarung besteht der Durchschnitt von $\dot{\mathfrak{K}}$ und $\mathfrak{F}$ aus einfachen geschlossenen, paarweise disjunkten Schnittlinien.

Wir betrachten den Durchschnitt von $\mathfrak{K}$ und $\mathfrak{F}$. Besteht er aus Elementarflächenstücken, so hat $\mathfrak{F}$ in $\mathfrak{K}$ die gewünschte Gestalt. Ist dies nicht der Fall, so gibt es Schnittlinien von $\dot{\mathfrak{K}}$ und $\mathfrak{F}$, die nicht Ränder von Elementarflächenstücken auf $\mathfrak{K} \cap \mathfrak{F}$ sind. Unter diesen Schnittlinien gibt es eine auf $\dot{\mathfrak{K}}$ innerste. Sie heiße s. Sie berandet auf $\dot{\mathfrak{K}}$ ein Elementarflächenstück $\mathfrak{e}$ derart, daß die weiteren Schnitte von $\mathfrak{e}$ und $\mathfrak{F}$ Ränder von Elementarflächenstücken in $\mathfrak{K} \cap \mathfrak{F}$ sind. Vermöge $\mathfrak{e}$ erkennt man, daß es in $\mathfrak{K}$ ein von s berandetes Elementarflächenstück $\mathfrak{g}$ gibt, für welches $\mathfrak{g} \cap \mathfrak{F} = \mathfrak{g} \cap \dot{\mathfrak{K}} = s$ ist. Wir schneiden $\mathfrak{F}$ längs s auf. In die beiden entstehenden Ränder fügen wir zwei Exemplare von $\mathfrak{g}$ ein und heben diese voneinander ab. (Fig. 8 zeigt den Sachverhalt schematisch im Schnitt.) Das Abheben werde ins Innere von $\mathfrak{K}$ vorgenommen, so daß sich $\mathfrak{F}$ außerhalb $\mathfrak{K}$ nicht ändert.

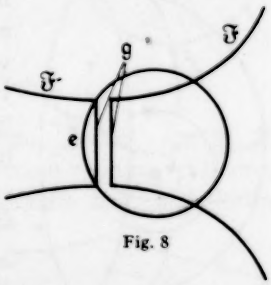

Fig. 8

Zerfällt $\mathfrak{F}$ bei dieser Operation (ξ-Operation bei HAKEN) nicht, so ist nachzuprüfen, ob $\mathfrak{F}$ weiterhin charakteristisch ist. Ist dies nicht mehr der Fall, so ist die Normalisierung gescheitert. Zerfällt $\mathfrak{F}$, so ist entweder eine der beiden entstehenden Flächen charakteristisch, so daß sie an die Stelle von $\mathfrak{F}$ treten kann, oder es scheitert die Normalisierung. Wir merken an: Tritt kein Zerfall ein, so erhöht sich $\chi(\mathfrak{F})$ um 2, tritt Zerfall ein, so ist die Charakteristik der neuen Fläche nicht kleiner als diejenige der alten.

Beispiel 1. Charakteristische Flächen seien 2-Sphären, die M nicht zerlegen. Durch die Operation zerfällt $\mathfrak{F}$ in zwei 2-Sphären, wovon mindestens eine charakteristisch ist. Durch eine solche ist $\mathfrak{F}$ zu ersetzen.

Beispiel 2. M möge durch jede 2-Sphäre zerlegt werden. $\dot{M}$ sei nicht zusammenhängend. Charakteristische Flächen seien 2-Sphären, die M so zerlegen, daß jeder der beiden Teile mindestens eine Komponente von $\dot{M}$ enthält. Bei der Operation zerfällt $\mathfrak{F}$ in zwei 2-Sphären, wovon mindestens eine charakteristisch ist. Zu einer solchen ist überzugehen.

Beispiel 3. C sei eine einfache geschlossene Kurve auf $\dot{M}$, die Unterkomplex von $\dot{M}$ bezüglich Σ ist und die auf $\dot{M}$ kein Elementarflächenstück berandet. Charakteristisch seien diejenigen von C berandeten, orientierbaren Flächen, die unter allen orientierbaren Flächen mit Rand C von maximaler Charakteristik sind. Bei der Operation zerfällt $\mathfrak{F}$ in eine 2-Sphäre und eine charakteristische Fläche, die an Stelle von $\mathfrak{F}$ tritt. Wegen der Gestalt von C waren hier übrigens der 3. und 5. Schritt gegenstandslos (vgl. die Anmerkung am Ende dieses Abschnitts 2.2).

Wir nehmen weiterhin an, daß die Normalisierung nicht scheitert, so daß durch Wiederholung der Operation erreicht wird, daß der Durchschnitt von $\mathfrak{F}$ mit allen Kugeln von Σ' aus Elementarflächenstücken besteht.

Waren nach dem 6. Schritt zylindrische Stücke von $\mathfrak{F}$ in den Balken vorhanden, so sind diese jetzt verschwunden, wenn nicht $\mathfrak{F}$ aus einem solchen zylindrischen Stück in einem Balken und zwei Elementarflächenstücken in den anstoßenden Kugeln besteht. Solche trivialen 2-Sphären werden jedoch unter den charakteristischen Flächen nie vorkommen.

8. Schritt. Weitere Vereinfachung von $\dot{\mathfrak{F}}$. Es sei $\mathfrak{K}$ eine Kugel von Σ', die an $\dot{M}$ anstößt (Fig. 9). Auf ihrem Rande $\dot{\mathfrak{K}}$ liegt ein Randsee $\mathfrak{r}$, der den Durchschnitt von $\mathfrak{K}$ und $\dot{M}$ ausmacht ($P\dot{M}$). $\mathfrak{i}$ sei eine an $\mathfrak{r}$ anstoßende Insel (PK) auf $\dot{\mathfrak{K}}$. Es kann sein, daß eine Schnittlinie von $\mathfrak{F}$ mit $\dot{\mathfrak{K}}$ auf $\mathfrak{i}$ eine Schleife bildet, die $\mathfrak{r}$ mit sich selbst verbindet. Solche Schleifen sollen beseitigt werden.

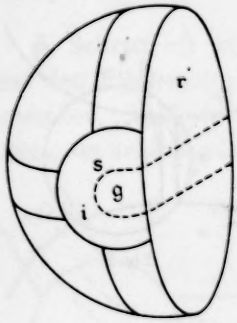

Fig. 9

In ihrer Umgebung hat $\mathfrak{F}$ tunnelartige Gestalt zu $\dot{M}$ hin. Sind Schleifen dieser Art auf $\mathfrak{i}$ vorhanden, so gibt es unter ihnen eine innerste, die s heiße. s schneidet von $\mathfrak{i}$ nach $\mathfrak{r}$ hin ein Elementarflächenstück $\mathfrak{g}$ ab, das von weiteren Schnitten mit $\mathfrak{F}$ frei ist. Wir schneiden $\mathfrak{F}$ längs s auf, fügen $\mathfrak{g}$ doppelt ein und heben die beiden Exemplare von $\mathfrak{g}$ nach beiden Seiten von $\mathfrak{i}$ so ab, daß s aus dem Durchschnitt von $\mathfrak{F}$ und $\mathfrak{i}$ verschwindet und keine neuen Schnitte entstehen (Rand-ξ-Operation). Danach bildet $\dot{\mathfrak{F}}$ eine Schleife auf dem $K\dot{M}$-Flächenstück, das den Durchschnitt von $\dot{M}$ mit dem in $\mathfrak{i}$ anstoßenden Balken ausmacht, und möglicherweise auch eine Schleife auf $\mathfrak{r}$.

Zerfällt $\mathfrak{F}$ bei dieser Operation nicht, so ist nachzuprüfen, ob $\mathfrak{F}$ weiterhin charakteristisch ist. Ist dies nicht der Fall, so ist die Normalisierung gescheitert. Zerfällt $\mathfrak{F}$, so ist entweder eine der entstehenden Flächen charakteristisch, so daß sie an Stelle von $\mathfrak{F}$ treten kann, oder es scheitert die Normalisierung. Zerfällt $\mathfrak{F}$ nicht, so erhöht sich $\chi(\mathfrak{F})$ um 1. Zerfällt $\mathfrak{F}$, so ist die Charakteristik der neuen Fläche nicht kleiner als diejenige der alten. Ein Beispiel wird sich in § 5 ergeben.

Wir nehmen weiterhin an, daß die Normalisierung nicht scheitert, so daß durch Wiederholung der Operation erreicht wird, daß Schleifen der betrachteten Art auf den Inseln nicht mehr vorhanden sind. Das Normalisierungsverfahren ist hierauf vom 5. Schritt an zu wiederholen, und zwar so oft wie nötig. Dies kann nur endlich viele Male geschehen, da sich $\dot{\mathfrak{F}}$ zunehmend vereinfacht. Die Anzahl der Querschnitte, die $\dot{\mathfrak{F}}$ auf den zu Σ' gehörigen Flächenstücken von $\dot{M}$ bildet, nimmt ab.

9. Schritt. Beseitigung von Falten. Es sei $\mathfrak{K}$ wieder eine Kugel von Σ' mit dem Rande $\dot{\mathfrak{K}}$. s sei eine Schnittlinie von $\mathfrak{F}$ und $\dot{\mathfrak{K}}$. s läuft auf $\dot{\mathfrak{K}}$ über Inseln und Brücken, möglicherweise auch über den Randsee von $\dot{\mathfrak{K}}$, falls $\mathfrak{K}$ an $\dot{M}$ anstößt. Wegen des 4. Schrittes bildet s auf den Brücken Schnitte, die verschiedene anstoßende Inseln verbinden (vgl. Fig. 5 sowie Fig. 2 und 4 PF). Nehmen wir nun an, daß s mehr als einmal über dieselbe Brücke $\mathfrak{b}$

führt. Es gibt dann auf $\mathfrak{b}$ ein innerstes Paar von Bögen auf s, etwa s_1 und s_2. Sie schneiden aus $\mathfrak{b}$ einen Streifen $\mathfrak{s}$ aus, der keinen weiteren Bogen von s enthält (Fig. 10). Da $\dot{\mathfrak{K}}$ von s in zwei Elementarflächenstücke zerlegt wird, können Bögen auf $\mathfrak{s}$, die von anderen Schnittlinien herrühren, nur in Paaren auftreten. Es kann daher angenommen werden, daß auf $\mathfrak{s}$ keine weiteren Schnittbögen von $\mathfrak{F}$ mit $\dot{\mathfrak{K}}$ liegen, weil man sonst statt s eine andere Schnittlinie betrachten könnte.

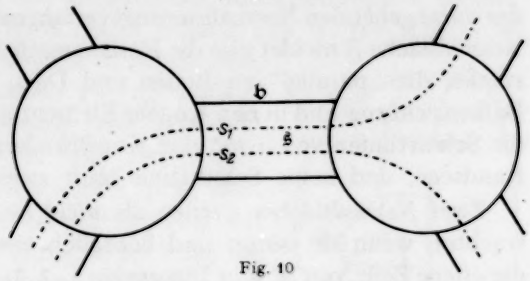

Fig. 10

An die Brücke $\mathfrak{b}$ stößt eine Platte $\mathfrak{P}$ an. In dieser Platte $\mathfrak{P}$ liegen zwei Stücke $\mathfrak{f}_1$ und $\mathfrak{f}_2$ von $\mathfrak{F}$, die in s_1 und s_2 an $\mathfrak{K}$ anstoßen (Fig. 11, $\mathfrak{K}$ und $\mathfrak{P}$ sind nicht mitgezeichnet). Wegen des 7. Schrittes berandet s auf $\mathfrak{F}$ ein Elementarflächenstück $\mathfrak{e}$, das in $\mathfrak{K}$ liegt. $\mathfrak{e}$ verbindet $\mathfrak{f}_1$ und $\mathfrak{f}_2$ griffartig; $\mathfrak{f}_1$, $\mathfrak{f}_2$ und $\mathfrak{e}$ bilden eine „Falte" von $\mathfrak{F}$. $\mathfrak{K}$ läßt sich durch ein Elementarflächenstück $\mathfrak{g}$ von $\mathfrak{e}$ her nach $\mathfrak{b}$ hin aufschlitzen, wobei $\mathfrak{g}$ mit $\mathfrak{F}$ nur einen auf $\mathfrak{e}$ liegenden Randbogen gemein hat. Man kann daher $\mathfrak{F}$ längs $\mathfrak{g}$ in $\mathfrak{P}$ hinein einbeulen und danach (vgl. 4. Schritt) $\mathfrak{f}_1$ und $\mathfrak{f}_2$ aus dem Durchschnitt von $\mathfrak{F}$ und $\mathfrak{P}$ verdrängen, während sich an dem sonstigen Verlauf von $\mathfrak{F}$ in $\mathfrak{P}$ und den übrigen Platten nichts ändert. Auch an $\dot{\mathfrak{F}}$ ändert sich nichts.

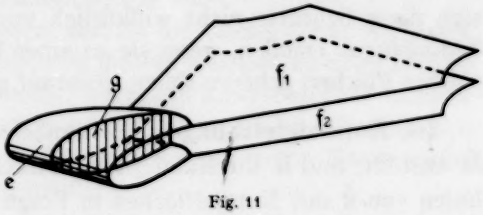

Fig. 11

Das Normalisierungsverfahren ist nun vom 6. Schritt an zu wiederholen, und zwar so lange, bis dieser 9. Schritt gegenstandslos wird, also keine Schnittlinie von $\mathfrak{F}$ mit einem Kugelrande zweimal über dieselbe Brücke führt. Da jedesmal die Anzahl der Stücke von $\mathfrak{F}$, die in den Platten liegen, um 2 abnimmt, sind nur endlich viele Wiederholungen möglich. Danach sind für die Schnittlinien von $\mathfrak{F}$ mit den Kugelrändern auch keine Schleifen auf den Inseln zu Brücken hin vorhanden.

Das Normalisierungsverfahren ist dann beendet und die gewünschte Gestalt von $\mathfrak{F}$ erreicht.

Anmerkung. Ist $\dot{\mathfrak{F}}$ fest vorgegeben, so sind der 3., 5. und 8. Schritt nicht erlaubt. Diese Schritte sind jedoch gegenstandslos, wenn $\dot{\mathfrak{F}}$ als Unterkomplex von $\dot{M}$ bezüglich der ursprünglichen Zellenzerlegung Σ gegeben ist. $\dot{\mathfrak{F}}$ meidet dann die Randseen (3. Schritt) und bildet auf jedem anderen Flächenstück von $\dot{M}$ bezüglich Σ' höchstens einen Querschnitt. Daher können weder Schleifen von $\dot{\mathfrak{F}}$ (5. Schritt) noch Schleifen nach Art der Fig. 9 (8. Schritt) vorhanden sein.

Es bedeutet übrigens keine Einschränkung der Allgemeinheit, bei fest vorgegebenem $\dot{\mathfrak{F}}$ anzunehmen, daß $\dot{\mathfrak{F}}$ Unterkomplex von $\dot{M}$ bezüglich Σ ist. Ist dies noch nicht der Fall, so ist Σ durch eine geeignete (explizit anzugebende) Unterteilung zu ersetzen.

2.3. Normalflächen. Es sei jetzt $\mathfrak{F}$ eine beliebige Fläche auf M, deren Rand $\dot{\mathfrak{F}}$ auf $\dot{M}$ liegt. $\mathfrak{F}$ heißt Normalfläche (bezüglich Σ'), wenn die Schritte des vorangehenden Normalisierungsverfahrens für $\mathfrak{F}$ gegenstandslos sind. Eine Normalfläche $\mathfrak{F}$ meidet also die Restraumstücke, bildet in den Platten Flächenstücke, die „parallel" zu Boden und Dach sind, in den Balken Bänder in Balkenrichtung und in den Kugeln Elementarflächenstücke. Außerdem bilden die Schnittlinien von $\mathfrak{F}$ mit den Kugelrändern keine Schleifen auf Inseln oder Randseen, und keine Schnittlinie läuft zweimal über dieselbe Brücke.

Zwei *Normalflächen* werden als *nicht wesentlich verschieden* (ähnlich) betrachtet, wenn sie isotop sind bezüglich einer Isotopie, bei der kein Punkt die offene Zelle von Σ' (der Dimension 1, 2, 3) verläßt, der er anfangs angehört. Eine solche *Isotopie* wird auch als *unwesentlich* bezeichnet.

Eine Normalfläche $\mathfrak{F}$ ist durch ihre Schnitte mit den Kugelrändern bis auf unwesentliche Isotopie bestimmt. Diese Schnitte legen bereits fest, wie $\mathfrak{F}$ in den Kugeln, Balken und Platten zu verlaufen hat. Natürlich lassen sich diese Schnitte nicht willkürlich vorgeben, sie müssen Verträglichkeitsbedingungen erfüllen, wenn sie zu einer Normalfläche oder zu einem System solcher Flächen gehören sollen. Hierauf gehen wir jetzt ein.

2.4. Normalgleichungen. Sei zunächst $\mathfrak{K}$ eine Kugel von Σ', die nicht an $\dot{M}$ anstößt, und $\dot{\mathfrak{K}}$ ihr Rand. Wir betrachten auf $\dot{\mathfrak{K}}$ Kurven, die als Schnittlinien von $\dot{\mathfrak{K}}$ mit Normalflächen in Frage kommen, also einfache geschlossene Kurven, welche Brücken und Inseln überqueren, die Seen meiden und keine Brücke zweimal überkreuzen. „Überqueren" soll besagen, daß keine Schleifen vorhanden sind und daß die Kurve nicht auf einer einzigen Insel oder Brücke liegt. Bezüglich unwesentlicher Isotopie gibt es nur endlich viele Klassen solcher Kurven. Wir bezeichnen diese Klassen als *Schnitt-Typen* auf $\dot{\mathfrak{K}}$. Sei nun $\mathfrak{K}$ eine Kugel, die an $\dot{M}$ anstößt. $\dot{\mathfrak{K}}$ sei wieder der Rand von $\mathfrak{K}$. Auf $\dot{\mathfrak{K}}$ liegt ein Randsee, der den Durchschnitt von $\mathfrak{K}$ und $\dot{M}$ ausmacht. Die entsprechenden Kurven können jetzt außer Inseln und Brücken auch den Randsee überqueren, wobei zwischen je zwei Überquerungen des Randsees mindestens eine Brückenüberquerung liegt. Daher gibt es auch jetzt nur endlich viele *Schnitt-Typen* auf $\dot{\mathfrak{K}}$.

Weil M kompakt ist, gibt es für alle Kugelränder insgesamt nur endlich viele Schnitt-Typen. Sie werden auf irgendeine, aber weiterhin feste Weise numeriert und mit $\mathfrak{P}_1, \mathfrak{P}_2, \ldots, \mathfrak{P}_n$ bezeichnet. Die Schnitt-Typen lassen sich aus dem Schema von Σ' bestimmen.

Will man zwei oder mehrere Schnitt-Typen, die zum selben Kugelrand gehören, durch repräsentierende Kurven darstellen, so kann es sein, daß sich Kreuzungen der Repräsentanten nicht vermeiden lassen. Ein Teilsystem des

Systems aller Schnitt-Typen heißt *realisierbar* (verträglich), wenn es durch paarweise disjunkte Kurven auf den Kugelrändern dargestellt werden kann. Unter einer *Realisierung* ist stets eine solche Darstellung zu verstehen. Von einem vorgegebenen Teilsystem der Schnitt-Typen ist finit entscheidbar, ob es realisierbar ist oder nicht. Man erkennt dies, wenn man für jeden Kugelrand die darzustellenden Schnitt-Typen der Reihe nach zu realisieren versucht und jeweils die Zerlegung des Kugelrandes durch die bereits vorhandenen Kurven berücksichtigt. Wird jedem Schnitt-Typ $\mathfrak{P}_i$ $(i=1, 2, \ldots, n)$ eine nicht-negative, ganze Vielfachheit x_i zugeordnet, so sagen wir, daß ein *Schnittschema* $\mathfrak{x}=\{x_i\}$ vorliegt (P-Zahlensystem). Eine *Realisierung* von $\mathfrak{x}$ besteht aus paarweise disjunkten Kurven auf den Kugelrändern, so daß jeweils x_i Kurven vom Typ $\mathfrak{P}_i$ sind. Man erkennt leicht, daß $\mathfrak{x}$ genau dann realisierbar ist, wenn es das System derjenigen Schnitt-Typen ist, die in $\mathfrak{x}$ eine positive Vielfachheit besitzen.

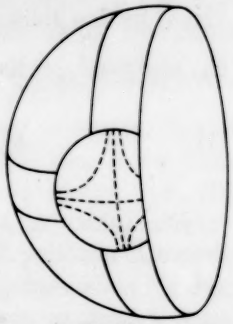

Fig. 12

Eine Normalfläche $\mathfrak{F}$ bestimmt ein realisierbares Schnittschema. Die Umkehrung gilt nicht; der Verlauf von $\mathfrak{F}$ in den Balken gibt eine weitere Bedingung. Dazu betrachten wir auf den Inseln Querschnitte, die verschiedene anstoßende Brücken verbinden, und für die Inseln, die an Randseen anstoßen, auch solche Querschnitte, die von diesem Randsee zu einer Brücke führen (Fig. 12, Querschnitte gestrichelt). Bezüglich unwesentlicher Isotopie gibt es insgesamt nur endlich viele Klassen solcher Querschnitte. Diese Klassen mögen *Bogentypen* heißen. Sie seien weiterhin mit einer festen Numerierung versehen und mit $\mathfrak{K}_1, \mathfrak{K}_2, \ldots, \mathfrak{K}_m$ bezeichnet.

Sei s eine Kurve auf einem Kugelrande, die einen Schnitt-Typ $\mathfrak{P}_i$ darstellt. s bildet auf den Inseln Querschnitte, die Bogentypen darstellen, wobei keine zwei solchen Querschnitte vom gleichen Bogentyp sind. Von den Bogentypen, die vermöge s erhalten werden, sagen wir, daß sie im Schnitt-Typ $\mathfrak{P}_i$ vorhanden sind. Sei

$$(1) \qquad a_{ji} = \begin{cases} 1, & \text{wenn der Bogentyp } \mathfrak{K}_j \text{ im Schnitt-Typ } \mathfrak{P}_i \text{ vorhanden ist,} \\ 0 & \text{sonst.} \end{cases}$$

Vermöge der Balken ergibt sich eine paarweise Zuordnung von Bogentypen. Diese Zuordnung entsteht, anschaulich gesprochen, durch „Translation in Balkenrichtung" (Fig. 13). Seien $\mathfrak{i}_1$ und $\mathfrak{i}_2$ die beiden Inseln an den Enden eines Balkens $\mathfrak{B}$. b_1 bzw. b_2 sei ein Querschnitt auf $\mathfrak{i}_1$ bzw. $\mathfrak{i}_2$. b_1 stelle den Bogentyp $\mathfrak{K}_j$, b_2 den Bogentyp $\mathfrak{K}_k$ dar. $\mathfrak{K}_j$ und $\mathfrak{K}_k$ heißen *gekoppelt*, wenn sich die Endpunkte von b_1 und b_2 auf dem Rande von $\mathfrak{B}$ durch zwei Querschnitte auf Streifen verbinden lassen, in denen Platten bzw. $\dot{M}$ an $\mathfrak{B}$

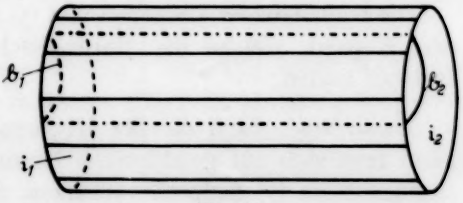

Fig. 13

anstoßen, wenn also b_1 und b_2 zum Rande eines Bandes gehören, das auf einer Normalfläche liegen kann. Sei

$$(2) \qquad b_{jh} = \begin{cases} 1, & \text{wenn die Bogentypen } \mathfrak{K}_j \text{ und } \mathfrak{K}_h \text{ gekoppelt sind,} \\ 0 & \text{sonst.} \end{cases}$$

Ist nun $\mathfrak{x} = \{x_i\}$ das Schnittschema einer Normalfläche $\mathfrak{F}$, so schneidet $\mathfrak{F}$ die Inseln in $\sum_{i=1}^{n} a_{ji} x_i$ Querschnitten vom Bogentyp $\mathfrak{K}_j$. Aus dem Verlauf von $\mathfrak{F}$ in den Balken folgt $\sum_{i=1}^{n} a_{ji} x_i = \sum_{i=1}^{n} a_{hi} x_i$ für gekoppelte Bogentypen $\mathfrak{K}_j$, $\mathfrak{K}_h$, also für $b_{jh} = 1$. Es gilt also

$$(3) \qquad b_{jh} \sum_{i=1}^{n} (a_{ji} - a_{hi}) x_i = 0 \qquad j, h = 1, 2, \ldots, m.$$

Die Zahlen a_{ji}, b_{jh} sind durch das Schema von Σ' bestimmt. (3) heiße das zu Σ' gehörige *Normalgleichungssystem*. Eine nicht-negative, ganze Lösung von (3) heiße *realisierbar*, wenn sie ein realisierbares Schnittschema ist. Hiernach gehört zu einer Normalfläche und allgemeiner zu einem System von (paarweise disjunkten) Normalflächen eine realisierbare Lösung von (3). Ist umgekehrt $\mathfrak{x}$ eine realisierbare Lösung von (3), so ist $\mathfrak{x}$ das Schnittschema eines Normalflächensystems $\mathfrak{F}$, und es ist $\mathfrak{F}$ bis auf unwesentliche Isotopie eindeutig bestimmt. Man kann nämlich zunächst die Schnitte von $\mathfrak{F}$ mit den Kugelrändern angeben und in diese Schnitte paarweise disjunkte Elementarflächenstücke innerhalb der Kugeln einspannen, weil $\mathfrak{x}$ realisierbar ist. Danach lassen sich die Bänder anfügen, die $\mathfrak{F}$ in den Balken bildet. Daß $\mathfrak{x}$ Lösung von (3) ist, besagt nämlich, daß für jeden Balken die an beiden Enden vorhandenen Bogensysteme auseinander durch „Translation in Balkenrichtung" entstehen. Der bisher aufgebaute Teil von $\mathfrak{F}$ wird berandet von einfachen geschlossenen, paarweise disjunkten Kurven auf $\dot{M}$ und den Plattenrändern, wobei keine Kurve über Boden oder Dach einer Platte führt. Daher kann $\mathfrak{F}$ noch in den Platten ergänzt werden. Aus dem Aufbau geht hervor, daß $\mathfrak{F}$ aus (paarweise disjunkten) Normalflächen besteht, und dieser Aufbau ist bis auf unwesentliche Isotopie der einzig mögliche.

Die realisierbaren Lösungen von (3) entsprechen also umkehrbar eindeutig den Klassen, welche die Normalflächensysteme bezüglich unwesentlicher Isotopie bilden.

Anmerkung. Sind die für M charakteristischen Flächen geschlossen, so kann man sich auf geschlossene Normalflächen beschränken und das Gleichungssystem (3) dadurch verkürzen, daß man alle Schnitt-Typen wegläßt, die Randseen kreuzen. Ist der Rand charakteristischer Flächen als Unterkomplex von $\dot{M}$ bezüglich Σ vorgegeben, so kann man sich auf Normalflächen beschränken, deren Randkomponenten zu Komponenten des vorgegebenen Randes unwesentlich isotop sind, womit sich (3) ebenfalls reduzieren läßt.

2.5. Fundamentalflächen. Nennen wir eine Lösung von (3) zulässig, wenn sie ganz, nicht-negativ und nicht-trivial ist. Eine Linearkombination zulässiger Lösungen heiße zulässig, wenn ihre Koeffizienten ganz, nicht-negativ und nicht sämtlich null sind. Es gibt ein eindeutig bestimmtes, endliches System Ξ zulässiger Lösungen von (3) derart, daß jede zulässige Lösung von (3) eine zulässige Linearkombination von Lösungen aus Ξ ist und daß keine Lösung aus Ξ Summe zweier zulässiger Lösungen ist. Die zu Ξ gehörigen Lösungen mögen *Fundamentallösungen* von (3) heißen.

Zur Bestimmung von Ξ bemerken wir zunächst, daß durch (3) und $x_i \geqq 0$ für $i = 1, 2, \ldots, n$ ein konvexer Kegel K im R^n definiert wird. Der Durchschnitt von K mit der Hyperebene $\sum_{i=1}^{n} x_i = 1$ ist ein kompakter konvexer Bereich B. Die Ecken von B sind rational (Punkte mit rationalen Koordinaten). Die rationalen Punkte aus B sind Linearkombinationen der Ecken von B mit nicht-negativen, rationalen Koeffizienten, wobei sich jede Ecke nur auf eine Weise, nämlich durch sich selbst, darstellen läßt. Die rationalen Punkte von K sind rationale Vielfache von rationalen Punkten aus B und daher nicht-negative, rationale Linearkombinationen der Ecken von B, wobei sich die Punkte auf einer Kante von K nur als Vielfache der entsprechenden Ecke von B darstellen lassen.

Für eine zulässige Lösung $\mathfrak{x} = \{x_i\}$ von (3) sei nun $|\mathfrak{x}| = \sum_{i=1}^{n} x_i$ die Norm. Auf jeder Kante von K liegt eine zulässige Lösung kleinster Norm. Sie muß jedenfalls zu Ξ gehören. Seien $\mathfrak{x}_1, \mathfrak{x}_2, \ldots, \mathfrak{x}_t$ die so durch die Kanten von K bestimmten Lösungen aus Ξ. Jede zulässige Lösung $\mathfrak{x}$ von (3) ist eine Linearkombination dieser Lösungen mit nicht-negativen, rationalen Koeffizienten. Ist dabei $|\mathfrak{x}| > \sum_{\tau=1}^{t} |\mathfrak{x}_\tau|$, so ist wegen der Linearität der Norm mindestens ein Koeffizient größer als 1, etwa derjenige von $\mathfrak{x}_\varrho$ $(1 \leqq \varrho \leqq t)$. Dann ist $\mathfrak{x} - \mathfrak{x}_\varrho$ eine zulässige Lösung von (3), und wegen $\mathfrak{x} = (\mathfrak{x} - \mathfrak{x}_\varrho) + \mathfrak{x}_\varrho$ gehört $\mathfrak{x}$ sicher nicht zu Ξ. Zu $\mathfrak{x}_1, \mathfrak{x}_2, \ldots, \mathfrak{x}_t$ nehme man nun zunächst alle zulässigen Lösungen von (3) hinzu, deren Norm nicht größer als $\sum_{\tau=1}^{t} |\mathfrak{x}_\tau|$ ist, und lasse danach alle diejenigen wieder weg, die Summen oder echte Vielfache von anderen sind. Damit erhält man Ξ.

Zu einer realisierbaren Fundamentallösung $\mathfrak{x}$ von (3) gehört eine Normalfläche $\mathfrak{F}$. Bestände nämlich $\mathfrak{F}$ aus mehreren Flächen, so wäre $\mathfrak{x}$ die Summe der Schnittschemata dieser Flächen, also nicht fundamental. Die zu den realisierbaren Fundamentallösungen von (3) gehörigen Normalflächen mögen *Fundamentalflächen* heißen.

2.6. Reduktion. Fassen wir zunächst zusammen: Aus der vorgegebenen Zellenzerlegung Σ der Mannigfaltigkeit M kann die zugehörige Normalzerlegung Σ' hergestellt werden und damit auch das Normalgleichungssystem

(3). Die Fundamentallösungen von (3) lassen sich berechnen, und ihre Realisierbarkeit ist auf endliche Weise entscheidbar. Zu jeder realisierbaren Fundamentallösung läßt sich eine Fundamentalfläche herstellen. Damit wird ein Vorrat von endlich vielen Flächen auf endliche Weise aufgewiesen. Einen solchen Vorrat wollen wir *repräsentativ* nennen.

Es wird sich nun darum handeln, die Frage nach charakteristischen Flächen auf die Untersuchung eines repräsentativen Vorrats von Fundamentalflächen zu reduzieren.

Dabei ist zunächst eine Bemerkung über den Rand der charakteristischen Flächen erforderlich. Es kann sein, daß dieser Rand ein Kurvensystem C auf $\dot{M}$ ist, das als Unterkomplex von $\dot{M}$ bezüglich Σ fest vorgegeben ist. In diesem Falle vereinbaren wir, daß ein repräsentativer Vorrat von Fundamentalflächen noch folgender Bedingung genügen soll: Ist für eine Fläche des Vorrats der Rand unwesentlich isotop zu C, so stimmt er sogar mit C überein. Durch diese Bedingung wird die Aufweisbarkeit des Vorrats offenbar nicht zerstört.

Für die angekündigte Reduktion ist ein Kompliziertheitsbegriff erforderlich. Unter dem *Plattengrad* (F-Zahl) einer (beliebigen) Normalfläche $\mathfrak{F}$ verstehen wir die Anzahl der Elementarflächenstücke, die den Durchschnitt von $\mathfrak{F}$ mit den Platten von Σ' ausmachen. Sind $\mathfrak{F}$ und $\mathfrak{G}$ Normalflächen, so heißt $\mathfrak{F}$ *einfacher* als $\mathfrak{G}$, wenn der Plattengrad von $\mathfrak{F}$ kleiner als derjenige von $\mathfrak{G}$ ist.

Wir wollen nun annehmen, daß sich die folgenden drei Aussagen beweisen lassen:

I (Normalisierung). Charakteristische Flächen lassen sich normalisieren.

II (Reduktion). Wenn charakteristische Normalflächen existieren, so gibt es zu jeder nicht-fundamentalen charakteristischen Normalfläche eine einfachere charakteristische Normalfläche.

III (Signifikanz). Für jede Fläche eines repräsentativen Vorrats von Fundamentalflächen läßt sich auf endliche Weise feststellen, ob sie charakteristisch ist oder nicht.

Dann liegt ein endliches Verfahren vor, charakteristische Flächen aufzuweisen, wenn solche existieren. I und II besagen zusammen, daß charakteristische Fundamentalflächen vorhanden sein müssen, wenn es überhaupt charakteristische Flächen gibt. Es führt nämlich II nach endlich vielen Wiederholungen auf eine charakteristische Fundamentalfläche, weil der Plattengrad nicht unbeschränkt abnehmen kann. III ergibt dann ein endliches Verfahren, das entweder charakteristische Flächen aufweist oder die Existenz solcher Flächen verneint.

Der Beweis der Aussagen I, II, III ist von Anwendung zu Anwendung verschieden, weil die Definition der charakteristischen Flächen eingeht. Für die Normalisierung (I) wurde eine allgemeine Methode in 2.2 angegeben, für die jeweils noch die Ausführbarkeit der einzelnen Schritte nachzuprüfen ist.

Die Reduktion (II) beruht auf einem allgemeinen Ansatz, dem wir uns jetzt zuwenden. Zur Signifikanz (III) werden allgemeine Bemerkungen nicht benötigt.

§ 3. Erzeugung von Normalflächen durch Flächenpaare

3.1. Vorbemerkung. Wir werden zeigen, daß man eine (beliebige) Normalfläche $\mathfrak{F}$, die nicht Fundamentalfläche ist, aus zwei einfacheren, sich schneidenden Normalflächen $\mathfrak{F}_1$, $\mathfrak{F}_2$ durch Umschalten der Schnittlinien erhalten kann. Ist $\mathfrak{F}$ charakteristisch, so braucht keine der beiden Flächen $\mathfrak{F}_1$ und $\mathfrak{F}_2$ charakteristisch zu sein. Es ergibt sich aber zu $\mathfrak{F}$ eine einfachere charakteristische Normalfläche möglicherweise dadurch, daß man Schnittlinien in anderer (irregulärer) Weise umschaltet oder Teile von $\mathfrak{F}_1$ und $\mathfrak{F}_2$ wegläßt und die danach vorhandene Fläche normalisiert.

3.2. Paare von Flächensystemen zu einer Normalfläche. Es sei $\mathfrak{F}$ eine Normalfläche, die nicht fundamental ist, und $\mathfrak{x}$ ihr Schnittschema. $\mathfrak{x}$ ist als Summe zweier zulässiger Lösungen des Normalgleichungssystems darstellbar. Im allgemeinen ist dies auf mannigfache Weise möglich. Irgendeine solche Darstellung $\mathfrak{x}=\mathfrak{x}'+\mathfrak{x}''$ werde weiterhin betrachtet.

$\mathfrak{x}'$ und $\mathfrak{x}''$ sind realisierbar. Aus den Schnitten von $\mathfrak{F}$ mit den Kugelrändern (von Σ') erhält man eine Realisierung von $\mathfrak{x}'$ bzw. $\mathfrak{x}''$ dadurch, daß man einen Teil der Schnittlinien wegläßt. Gemäß 2.4 sind $\mathfrak{x}'$ und $\mathfrak{x}''$ Schnittschemata von Flächensystemen $\mathfrak{F}'$ und $\mathfrak{F}''$, die aus Normalflächen bestehen. $\mathfrak{F}'$ und $\mathfrak{F}''$ müssen sich schneiden, weil sonst $\mathfrak{F}'$ und $\mathfrak{F}''$ zusammen ein System von Normalflächen bildeten, welches dasselbe Schnittschema wie die Fläche $\mathfrak{F}$ hätte, aber aus mindestens zwei Flächen bestände, was wegen 2.4 unmöglich ist. Wir werden $\mathfrak{F}'$ und $\mathfrak{F}''$ so wählen, daß die Schnitte in übersichtlicher Lage sind.

Wir betrachten zunächst die Schnitte von $\mathfrak{F}$ mit den Kugelrändern. Entsprechend der Darstellung $\mathfrak{x}=\mathfrak{x}'+\mathfrak{x}''$ werden diese Kurven auf Realisierungen von $\mathfrak{x}'$ und $\mathfrak{x}''$ verteilt. Falls ein Schnitt-Typ $\mathfrak{P}_i$ in $\mathfrak{x}'$ und in $\mathfrak{x}''$ mit positiver Vielfachheit auftritt, darf diese Verteilung für die Kurven vom Typ $\mathfrak{P}_i$ in willkürlicher Weise vorgenommen werden; nur die richtige Vielfachheit für $\mathfrak{x}'$ und $\mathfrak{x}''$ ist wesentlich. Die Flächenstücke, die $\mathfrak{F}$ in den Kugeln bildet, werden entsprechend auf $\mathfrak{F}'$ und $\mathfrak{F}''$ verteilt.

Betrachten wir nun die Bänder, die $\mathfrak{F}$ in einem Balken $\mathfrak{B}$ bildet. Ihre Enden sind Schnittbögen von $\mathfrak{F}$ mit den beiden Inseln im Rande von $\mathfrak{B}$ (vgl. Fig. 7 und 13). Dabei stellen die Enden eines jeden Bandes ein Paar gekoppelter Bogentypen dar. Sind zu einem Paar gekoppelter Bogentypen mehrere Bänder von $\mathfrak{F}$ vorhanden, so schneiden diese aus $\mathfrak{B}$ ein „Brett" $\mathfrak{A}$ aus, das aus diesen Bändern und ihren Zwischenräumen besteht. $\mathfrak{A}$ trifft die übrigen Bänder von $\mathfrak{F}$ auf $\mathfrak{B}$ und die ihnen entsprechenden Bretter nicht. An den Enden von $\mathfrak{A}$ sind Schnittbögen von $\mathfrak{F}'$ vorgegeben, wobei die Anzahl an beiden Enden dieselbe ist, weil $\mathfrak{x}'$ Lösung von (3) ist. Dasselbe gilt für $\mathfrak{F}''$. Man erkennt nun ohne weiteres, daß sich die entsprechenden Bänder von $\mathfrak{F}'$

und $\mathfrak{F}''$ so wählen lassen, daß sie auf $\mathfrak{A}$ verlaufen und daß sich je zwei Bänder in höchstens einem Querschnitt schneiden. (Fig. 14 zeigt den Sachverhalt im Schnitt.) Die Bänder von $\mathfrak{F}'$ schneiden sich dabei nicht untereinander, gleiches gilt für die Bänder von $\mathfrak{F}''$. Damit ergibt sich, daß sich die Bänder von $\mathfrak{F}'$ und $\mathfrak{F}''$ in allen Balken so bestimmen lassen, daß sich je zwei Bänder in höchstens einem Querschnitt schneiden und daß sich nur solche Bänder schneiden, die zum selben Paar gekoppelter Bogentypen gehören.

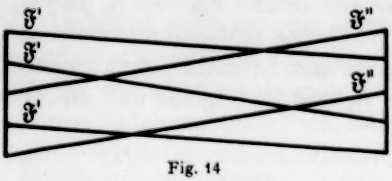

Fig. 14

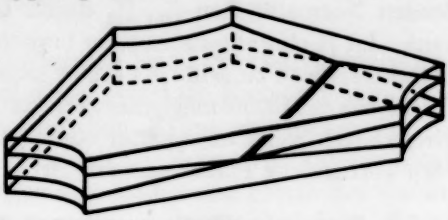
Fig. 15

$\mathfrak{F}'$ und $\mathfrak{F}''$ sind nun noch in den Platten durch Elementarflächenstücke zu ergänzen, deren Rand bereits festliegt. Dies kann offenbar so geschehen, daß sich diese Elementarflächenstücke nur in Querschnitten schneiden (Fig. 15). Kein solcher Querschnitt kann einen Balkenanstoß mit sich selbst verbinden, weil sich sonst in diesem Balken zwei Bänder mehrfach schneiden müßten. Jedoch können sich zwei Elementarflächenstücke in mehreren Querschnitten durchsetzen (wenn mehr als drei Balken anstoßen).

Damit sind $\mathfrak{F}'$ und $\mathfrak{F}''$ so bestimmt, daß der Durchschnitt von $\mathfrak{F}'$ und $\mathfrak{F}''$ aus einfachen, paarweise disjunkten Schnittlinien besteht, die geschlossen sind oder in $\dot{M}$ entspringen und enden, wobei jede Schnittlinie abwechselnd durch Balken und Platten läuft.

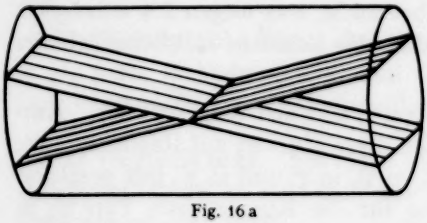
Fig. 16a

3.3. Umschalten von Schnittlinien. Die Schnitte von $\mathfrak{F}'$ und $\mathfrak{F}''$ sollen durch Umschalten beseitigt werden. Wir betrachten zunächst in einem Balken $\mathfrak{B}$

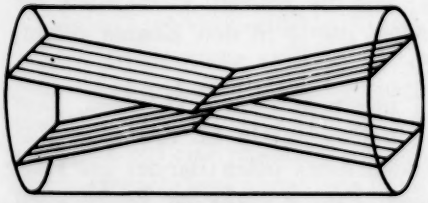
Fig. 16b

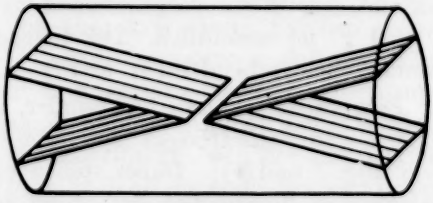
Fig. 16c

zwei sich schneidende Bänder $\mathfrak{a}'$ und $\mathfrak{a}''$ von $\mathfrak{F}'$ und $\mathfrak{F}''$ (Fig. 16a). Schneiden wir $\mathfrak{a}'$ und $\mathfrak{a}''$ längs ihres Schnittes auf, so lassen sich die vier entstehenden Elementarflächenstücke auf zwei verschiedene Weisen zu zwei disjunkten Elementarflächenstücken zusammenheften, nämlich entweder so, daß wieder zwei in Balkenrichtung verlaufende Bänder entstehen (Fig. 16b), oder so,

daß die beiden entstehenden Elementarflächenstücke faltenartig in $\mathfrak{B}$ liegen (Fig. 16c). Die Beseitigung des Schnittes von $\mathfrak{a}'$ und $\mathfrak{a}''$ gemäß Fig. 16b werde als *reguläre Umschaltung* bezeichnet, diejenige gemäß Fig. 16c als *irreguläre*. Setzt man die Umschaltung des Schnittes in eine anstoßende Platte fort (vgl. Fig. 15), so entstehen aus den beiden betroffenen Elementarflächenstücken zwei Elementarflächenstücke, die im Falle der regulären Umschaltung wieder „parallel" zu Boden und Dach der Platte liegen, und die im Falle der irregulären Umschaltung faltenartig liegen, so daß sie aus der Platte in die anstoßenden Balken und Kugeln verdrängt werden können. Setzt man nun die Umschaltung von der Platte noch weiter in einen anstoßenden Balken fort, so erhält man dort dieselbe Umschaltungsart, von der man ausgegangen war. Man bemerkt, daß eine kohärente Umschaltung einer Schnittlinie von $\mathfrak{F}'$ und $\mathfrak{F}''$ entweder überall regulär oder überall irregulär ist.

Für die geschlossenen Schnittlinien von $\mathfrak{F}'$ und $\mathfrak{F}''$ ergibt sich hieraus eine bemerkenswerte Folgerung, die später von Bedeutung sein wird: Ist eine geschlossene Schnittlinie s von $\mathfrak{F}'$ und $\mathfrak{F}''$ orientierungsumkehrend auf $\mathfrak{F}'$ (d. h. s besitzt ein Möbius-Band als Umgebung auf $\mathfrak{F}'$), so ist sie auch orientierungsumkehrend auf $\mathfrak{F}''$, weil sich andernfalls der Umschaltsinn längs s umkehrte.

Werden sämtliche Schnitte von $\mathfrak{F}'$ und $\mathfrak{F}''$ regulär umgeschaltet, so ergibt sich ein Flächensystem, das aus Normalflächen besteht. Es hat dasselbe Schnittschema wie $\mathfrak{F}$, weil an den Schnitten der Kugelränder mit $\mathfrak{F}'$ und $\mathfrak{F}''$ nichts geändert wird. Durch reguläres Umschalten aller Schnitte von $\mathfrak{F}'$ und $\mathfrak{F}''$ erhält man also $\mathfrak{F}$ (bis auf unwesentliche Isotopie).

3.4. Erzeugendes Normalflächenpaar. Wird nur ein Teil der Schnittlinien von $\mathfrak{F}'$ und $\mathfrak{F}''$ umgeschaltet (regulär oder irregulär), so können Flächensysteme entstehen, deren Flächen außer gegenseitigen Schnitten auch Selbstdurchdringungen (Singularitäten) besitzen. (Die Flächen können durch die Komponenten eines schnittfreien Urbildes erklärt werden.)

$\mathfrak{F}'$ und $\mathfrak{F}''$ bestehen zusammen aus mindestens zwei Flächen, während $\mathfrak{F}$ eine einzige Fläche ist. Weil $\mathfrak{F}$ aus $\mathfrak{F}'$ und $\mathfrak{F}''$ durch reguläres Umschalten aller Schnittlinien entsteht, muß es in dem System dieser Schnittlinien ein Teilsystem geben, das maximal bezüglich folgender Eigenschaft ist: Werden alle Schnitte des Teilsystems regulär umgeschaltet, so entstehen aus $\mathfrak{F}'$ und $\mathfrak{F}''$ genau zwei Flächen $\mathfrak{F}_1$ und $\mathfrak{F}_2$.

Aus der Forderung der Maximalität ergeben sich wichtige Folgerungen.

(1) $\mathfrak{F}_1$ und $\mathfrak{F}_2$ sind frei von Selbstdurchdringungen.

Besäße nämlich etwa $\mathfrak{F}_1$ eine Selbstdurchdringung, so könnte diese regulär umgeschaltet werden, und es wären danach weiterhin zwei oder drei Flächen vorhanden. Beides widerspricht der Maximalitätsforderung.

(2) $\mathfrak{F}_1$ und $\mathfrak{F}_2$ sind Normalflächen. Sind $\mathfrak{x}_1$ und $\mathfrak{x}_2$ die Schnittschemata von $\mathfrak{F}_1$ und $\mathfrak{F}_2$, so ist $\mathfrak{x} = \mathfrak{x}_1 + \mathfrak{x}_2$.

Dies folgt unmittelbar aus (1) und der Tatsache, daß nur reguläre Umschaltungen vorgenommen wurden.

(3) Der Plattengrad von $\mathfrak{F}$ ist die Summe der Plattengrade von $\mathfrak{F}_1$ und $\mathfrak{F}_2$.

Dies folgt daraus, daß $\mathfrak{F}$ aus $\mathfrak{F}_1$ und $\mathfrak{F}_2$ durch reguläres Umschalten entsteht.

(4) $\mathfrak{F}_1$ und $\mathfrak{F}_2$ schneiden sich. Werden irgendwelche Schnitte von $\mathfrak{F}_1$ und $\mathfrak{F}_2$ regulär umgeschaltet, so entsteht stets eine einzige Fläche (mit Selbstdurchdringungen, wenn nicht alle Schnitte umgeschaltet werden). $\mathfrak{F}$ entsteht aus $\mathfrak{F}_1$ und $\mathfrak{F}_2$ durch reguläres Umschalten aller Schnittlinien.

(5) Wird eine der beiden Flächen $\mathfrak{F}_1$ und $\mathfrak{F}_2$ durch eine Schnittlinie s zerlegt, so zerlegt s die andere Fläche nicht.

Man erhielte andernfalls vier Flächenteile, wenn man $\mathfrak{F}_1$ und $\mathfrak{F}_2$ längs s aufschneidet. Nach regulärer Umschaltung von s müßten wieder zwei Flächen vorhanden sein, was ausgeschlossen ist.

Außerdem gilt noch wie in 3.3

(6) Ist eine geschlossene Schnittlinie von $\mathfrak{F}_1$ und $\mathfrak{F}_2$ orientierungsumkehrend auf $\mathfrak{F}_1$, so ist sie auch orientierungsumkehrend auf $\mathfrak{F}_2$.

Schließlich gilt offenbar noch

(7) Werden Schnittlinien von $\mathfrak{F}_1$ und $\mathfrak{F}_2$ umgeschaltet, so ändert sich die Gesamtcharakteristik nicht. Insbesondere ist $\chi(\mathfrak{F}) = \chi(\mathfrak{F}_1) + \chi(\mathfrak{F}_2)$.

Zu $\mathfrak{F}$ werden statt $\mathfrak{F}'$ und $\mathfrak{F}''$ weiterhin $\mathfrak{F}_1$ und $\mathfrak{F}_2$ betrachtet. $\mathfrak{F}_1$ und $\mathfrak{F}_2$ heißen ein *erzeugendes Normalflächenpaar* für $\mathfrak{F}$.

3.5. Reduktionsmöglichkeiten. Es sei nun $\mathfrak{F}$ eine Normalfläche, die charakteristisch, aber nicht fundamental ist. $(\mathfrak{F}_1, \mathfrak{F}_2)$ sei ein erzeugendes Normalflächenpaar für $\mathfrak{F}$.

Ist $\mathfrak{F}_1$ oder $\mathfrak{F}_2$ charakteristisch, so hat man eine charakteristische Normalfläche vorliegen, die einfacher als $\mathfrak{F}$ ist.

Sei dies jetzt nicht der Fall. Schaltet man Schnittlinien von $\mathfrak{F}_1$ und $\mathfrak{F}_2$ (möglicherweise nicht alle) so um, daß mindestens eine Umschaltung irregulär erfolgt, so kann es sein, daß dabei eine Fläche $\tilde{\mathfrak{F}}$ entsteht, die frei von Selbstdurchdringungen und charakteristisch ist. Außer $\tilde{\mathfrak{F}}$ können noch andere Flächen entstehen, die Selbstdurchdringungen haben und $\tilde{\mathfrak{F}}$ schneiden dürfen. Betrachtet man die irreguläre Umschaltung eines Schnittbogens innerhalb einer Platte, so entstehen aus den beiden beteiligten Elementarflächenstücken von $\mathfrak{F}_1$ und $\mathfrak{F}_2$ zwei faltenartig liegende, wie in 3.3 bemerkt wurde. Diese können sich nicht gegenseitig, aber sich noch selbst und andere Flächenstücke schneiden. Durch die weiter vorzunehmenden Umschaltungen kann sich die faltenartige Lage noch komplizieren, und es können auch Flächenstücke entstehen, die mehrere Ränder besitzen. Es läßt sich aber jedenfalls die Platte

„senkrecht" so durchstechen, daß auf dem Durchstich die Anzahl der Schnittpunkte mit den jetzt vorliegenden Flächenstücken um mindestens zwei kleiner ist als die Anzahl der Flächenstücke von $\mathfrak{F}_1$ und $\mathfrak{F}_2$, die vor dem Umschalten in der Platte lagen. Hieraus folgt, daß man für die Normalisierung von $\tilde{\mathfrak{F}}$ gemäß 2.2 den 4. Schritt so vornehmen kann, daß die Anzahl der in den Platten liegenden Flächenstücke von $\tilde{\mathfrak{F}}$ kleiner als der Plattengrad von $\mathfrak{F}$ wird. Die ersten drei Schritte des Normalisierens sind für $\tilde{\mathfrak{F}}$ gegenstandslos. Die auf den 4. Schritt folgenden Schritte können die Zahl der Flächenstücke von $\tilde{\mathfrak{F}}$ in den Platten nur noch verringern, so daß man aus $\tilde{\mathfrak{F}}$ durch Normalisieren[7]) eine charakteristische Normalfläche erhält, deren Plattengrad kleiner als derjenige von $\mathfrak{F}$ ist.

Wir weisen noch auf einen Spezialfall hin. Es mögen $\mathfrak{F}_1$ und $\mathfrak{F}_2$ längs sämtlicher Schnittlinien aufgeschnitten werden. Wenn es gelingt, aus den nun vorliegenden Flächenteilen eine charakteristische Fläche $\tilde{\mathfrak{F}}$ so aufzubauen, daß nicht alle Flächenteile verwendet werden, so ist $\tilde{\mathfrak{F}}$ entweder eine der beiden Flächen $\mathfrak{F}_1$, $\mathfrak{F}_2$, und damit eine einfachere charakteristische Normalfläche als $\mathfrak{F}$, oder es ist $\tilde{\mathfrak{F}}$ eine Fläche, die aus $\mathfrak{F}_1$ und $\mathfrak{F}_2$ durch Umschalten gewisser Schnittlinien entsteht. Ist das letzte der Fall, so ist mindestens eine Umschaltung irregulär, weil bei nur regulären Umschaltungen wegen 3.4, (4) keine Flächenteile abfallen können. Nach dem Vorangehenden läßt sich aus $\tilde{\mathfrak{F}}$ eine charakteristische Normalfläche erhalten, die einfacher als $\mathfrak{F}$ ist.

§ 4. Zerfallende Verkettungen

4.1. Satz. *Ist eine Verkettung V in der S^3 gegeben, so läßt sich entscheiden, ob V zerfällt*[8]).

Unter einer Verkettung V wird hier ein Polygonsystem verstanden. V zerfällt, wenn die S^3 durch eine zu V disjunkte 2-Sphäre so in zwei Kugeln zerlegt werden kann, daß jede der beiden Kugeln mindestens eine Komponente von V enthält.

Daß V in der S^3 gegeben ist, soll bedeuten, daß eine reguläre Zellenzerlegung der S^3 und V als Unterkomplex dieser Zellenzerlegung gegeben sind. Ist zunächst V als Verkettung im R^3 durch eine normierte Projektion oder durch ein Polygonsystem gegeben, so läßt sich erreichen, daß V als Verkettung in der S^3 gegeben ist. Zu gegebener normierter Projektion läßt sich nämlich jedenfalls ein Polygonsystem im R^3 finden, das diese Projektion besitzt. Liegt ein Polygonsystem V im R^3 vor, so ist es im Innern eines geeigneten 3-Simplexes E^3 des R^3 enthalten, und es kann eine reguläre Zellenzerlegung von E^3 angegeben werden, für welche V Unterkomplex ist, beispielsweise so: Es kann angenommen werden, daß auf V keine drei aufeinander folgenden Kanten in einer Ebene liegen. Zerlegt man E^3 durch alle Ebenen, die durch je zwei aufeinander folgende Kanten von V gehen, so entsteht eine Zellenzerlegung

[7]) Wir nehmen an, daß dies ausführbar ist.

[8]) Haken [1].

der gewünschten Art. Faßt man nun E^3 als ein 3-Simplex einer S^3 auf, die in fünf 3-Simplexe zerlegt ist[9]), so erhält man V als gegebene Verkettung in der S^3.

Nach PAPAKYRIAKOPOULOS [3] ist der Satz gleichwertig damit, daß von einer in der S^3 gegebenen Verkettung V entschieden werden kann, ob die Fundamentalgruppe des Komplements von V ein freies Produkt ist oder nicht.

4.2. Beweisansatz. V wird zunächst aus der S^3 ausgebohrt, d. h. es werden die Komponenten von V mit paarweise disjunkten Vollringen umgeben, welche die Komponenten als Seelen[10]) besitzen, und das Innere dieser Vollringe wird aus der S^3 entfernt. Solche Vollringe kann man beispielsweise aus „Kugeln" und „Balken" nach Art von 2.1 erhalten. Durch das Ausbohren von V entsteht aus der S^3 eine berandete Mannigfaltigkeit M. Vermöge der vorhandenen Zellenzerlegung der S^3 und des Randes von M ergibt sich eine Zellenzerlegung Σ von M. M ist in die S^3 eingebettet, was im folgenden benutzt wird. $\dot{M}$ bezeichnet wieder den Rand von M.

Jede 2-Sphäre in M zerlegt M, weil die S^3 durch jede 2-Sphäre zerlegt wird. Die Aussage, daß V zerfällt, ist gleichbedeutend mit der Aussage, daß M durch eine 2-Sphäre so zerlegt werden kann, daß auf jedem Teil mindestens eine Komponente von $\dot{M}$ liegt. Von einer solchen 2-Sphäre wollen wir sagen, daß sie $\dot{M}$ trennt. Für M charakteristische Flächen seien nun diejenigen 2-Sphären auf M, die $\dot{M}$ trennen. Der Satz ist damit auf eine Aussage über die für M charakteristischen Flächen zurückgeführt.

Nach 2.6 sind die Aussagen I, II und III (Normalisierung, Reduktion, Signifikanz) zu beweisen.

Die Signifikanz ergibt sich einfach. Für jede Normalfläche $\mathfrak{F}$ (bezüglich der zu Σ gehörigen Normalzerlegung Σ' von M) ist vermöge $\chi(\mathfrak{F})$ feststellbar, ob sie eine 2-Sphäre ist. Sei dies für $\mathfrak{F}$ der Fall. $\mathfrak{F}$ bewirkt eine Unterteilung von Σ', vermöge deren rein kombinatorisch entschieden werden kann, ob $\mathfrak{F}$ charakteristisch ist.

Die Normalisierung ergibt sich ohne weiteres aus 2.2. Die den Rand betreffenden Schritte sind gegenstandslos. Von den übrigen ist nur der 7. Schritt keine isotope Deformation. Er ist jedoch ausführbar nach dem dort angegebenen Beispiel 2.

Nur II ist noch zu beweisen.

4.3. Reduktion. Nehmen wir an, daß es eine charakteristische Normalfläche $\mathfrak{F}$ gibt, die nicht fundamental ist. $(\mathfrak{F}_1, \mathfrak{F}_2)$ sei ein erzeugendes Normalflächenpaar für $\mathfrak{F}$ gemäß 3.4. Wegen 3.4, (7) und $\chi(\mathfrak{F}) = 2$ muß eine der beiden Flächen $\mathfrak{F}_1$, $\mathfrak{F}_2$, etwa $\mathfrak{F}_1$, die Charakteristik 2 und die andere, also $\mathfrak{F}_2$, die Charakteristik 0 haben. $\chi(\mathfrak{F}_1) = \chi(\mathfrak{F}_2) = 1$ ist nämlich ausgeschlossen, weil sonst $\mathfrak{F}_1$ und $\mathfrak{F}_2$ (singularitätenfreie) projektive Ebenen in M und damit in der S^3 wären, was unmöglich ist. Entsprechend folgt, daß $\mathfrak{F}_2$ ein Torus und kein

[9]) Vgl. [5], [6].
[10]) Vgl. [6], § 6.

Kleinscher Schlauch ist. $\mathfrak{F}_1$ ist eine 2-Sphäre. Ist $\mathfrak{F}_1$ charakteristisch, so liegt eine charakteristische Normalfläche vor, die einfacher als $\mathfrak{F}$ ist.

Sei nun $\mathfrak{F}_1$ nicht charakteristisch. Die 2-Sphäre $\mathfrak{F}_1$ trenne also $\dot{M}$ nicht. Dies bedeutet, daß $\mathfrak{F}_1$ die S^3 so in zwei Kugeln $\mathfrak{M}_1$, $\mathfrak{M}_2$ zerlegt, daß die Verkettung V im Inneren einer dieser beiden Kugeln liegt, etwa im Inneren von $\mathfrak{M}_1$. $\mathfrak{F}_1$ wird durch jede Schnittlinie von $\mathfrak{F}_1$ und $\mathfrak{F}_2$ zerlegt. Wegen 3.4, (5) zerlegt keine solche Schnittlinie $\mathfrak{F}_2$. Es folgt, daß $\mathfrak{F}_2$ durch die Schnitte von $\mathfrak{F}_1$ und $\mathfrak{F}_2$ in Kreisringe zerlegt wird, die abwechselnd in $\mathfrak{M}_1$ und in $\mathfrak{M}_2$ liegen. Wird von den beiden Teilen, in welche die S^3 durch $\mathfrak{F}$ zerlegt wird, der eine schwarz, der andere weiß gefärbt, so ergibt sich eine entsprechende Schwarz-Weiß-Färbung für diejenigen Bereiche der S^3, die beim Zerschneiden der S^3 längs $\mathfrak{F}_1$ und $\mathfrak{F}_2$ entstehen. Weil $\mathfrak{F}$ charakteristisch ist, können nicht alle Komponenten von V in den schwarzen oder in den weißen Bereichen liegen. Hieraus folgt: $\mathfrak{M}_1$ wird durch einen Kreisring $\mathfrak{r}$ auf $\mathfrak{F}_2$ so zerlegt, daß beide Teile Komponenten von V enthalten. $\mathfrak{r}$ bildet zusammen mit zwei Elementarflächenstücken auf $\mathfrak{F}_1$ eine charakteristische 2-Sphäre $\tilde{\mathfrak{F}}$. Nach 3.5 läßt sich aus $\tilde{\mathfrak{F}}$ eine charakteristische Normalfläche erhalten, die einfacher als $\mathfrak{F}$ ist. Dies war zu zeigen.

4.4. Zusatz. Zerfällt die Verkettung V in die Verkettungen V_1 und V_2, so wollen wir sagen, daß V Summe von V_1 und V_2 ist. Wenn dabei V_1 oder V_2 eine zerfallende Verkettung ist, so gibt es noch andere Darstellungen von V als Summe zweier Verkettungen. Im Vorangehenden ist beim 7. Schritt des Normalisierens und bei der Reduktion möglicherweise von einer Summendarstellung von V zu einer anderen übergegangen worden. Hat man durch eine Fundamentalfläche eine Summendarstellung von V gefunden, so läßt sich das Verfahren auf die Summanden erneut anwenden und so oft wiederholen, bis man V als Summe von nicht zerfallenden Verkettungen erhält. Die Anzahl dieser Summanden ist offenbar nicht größer als die Anzahl der Komponenten von V. Man kann leicht zeigen, daß sich V und seine nicht zerfallenden Summanden gegenseitig umkehrbar eindeutig bis auf semilineare Äquivalenz bestimmen.

§ 5. Das Geschlecht von Knoten

5.1. Satz. *Es sei $\hat{M}$ eine kompakte, möglicherweise berandete, 3-dimensionale Mannigfaltigkeit. k sei eine einfache, geschlossene Kurve im Innern von $\hat{M}$, die Unterkomplex einer gegebenen regulären Zellenzerlegung von $\hat{M}$ ist. Es läßt sich entscheiden, ob k auf $\hat{M}$ eine orientierbare Fläche berandet. Ist dies der Fall, so läßt sich eine von k berandete Fläche maximaler Charakteristik aufweisen. Entsprechende Aussagen gelten, wenn orientierbare Flächen durch beliebige Flächen ersetzt werden.*

Ein Spezialfall des Satzes ist die Entscheidbarkeit der Frage, ob k ein Kreis, d.h. Rand eines Elementarflächenstückes auf $\hat{M}$ ist [11]).

[11]) Vgl. HAKEN [1].

Ist speziell $\hat{M} = S^3$, so ergibt sich, daß das Geschlecht eines Knotens (SEIFERT [7]) bestimmt werden kann [11]). Damit ist insbesondere für einen Knoten k in der S^3 entscheidbar, ob die Fundamentalgruppe seines Komplements frei zyklisch ist. Wegen des von PAPAKYRIAKOPOULOS [4] bewiesenen Dehnschen Lemmas ist dies genau dann der Fall, wenn k ein Kreis ist. Für die Vorgabe eines Knotens in der S^3 gelten die Bemerkungen in 4.1.

5.2. Beweisansatz. Da k Unterkomplex von $\hat{M}$ ist, läßt sich eine Umgebung $\mathfrak{V}$ von k angeben, die aus „Kugeln" und „Balken" nach Art von 2.1 besteht. $\mathfrak{T}$ sei der Rand von $\mathfrak{V}$. $\mathfrak{T}$ ist ein Torus oder ein Kleinscher Schlauch. M entstehe aus $\hat{M}$ dadurch, daß das Innere von $\mathfrak{V}$ weggelassen wird. Aus der vorhandenen Zellenzerlegung von $\hat{M}$ und aus $\mathfrak{T}$ erhält man eine reguläre Zellenzerlegung Σ von M. Σ' sei wieder die zugehörige Normalzerlegung von M.

Beim Bestimmen von $\mathfrak{V}$ kann innerhalb eines „Balkens" von $\mathfrak{V}$ eine Meridianfläche $\mathfrak{m}$ angegeben werden, die k in nur einem Punkte trifft. (Wird $\mathfrak{V}$ längs $\mathfrak{m}$ aufgeschnitten, so entstehen aus $\mathfrak{V}$ und k eine Kugel mit unverknoteter Sehne.) Der Rand von $\mathfrak{m}$ sei m. Er ist ein Meridian von $\mathfrak{T}$. Die Meridianfläche $\mathfrak{m}$ kann so angegeben werden, daß m die Randseen auf $\mathfrak{T}$, in denen Kugeln von Σ' anstoßen, nicht trifft und daß m die übrigen Flächenstücke von Σ' auf $\mathfrak{T}$ in jeweils höchstens einem Querschnitt überquert.

Ist $\hat{\mathfrak{F}}$ eine Fläche auf $\hat{M}$, die k als Rand besitzt, so kann sie bei festem Rand isotop so deformiert werden, daß ihr Durchschnitt mit $\mathfrak{V}$ aus einem Kreisring und ihr Durchschnitt mit $\mathfrak{m}$ aus einem Querschnitt dieses Kreisringes besteht. Der außerhalb $\mathfrak{V}$ liegende Teil von $\hat{\mathfrak{F}}$ ist dann eine Fläche $\mathfrak{F}$, deren Rand auf $\mathfrak{T}$ liegt und m in nur einem Punkte schneidet. Liegt umgekehrt eine solche Fläche $\mathfrak{F}$ vor, so kann sie durch einen Kreisring auf $\mathfrak{V}$ zu einer von k berandeten Fläche $\hat{\mathfrak{F}}$ ergänzt werden [12]). $\mathfrak{F}$ und $\hat{\mathfrak{F}}$ stimmen im Orientierbarkeitscharakter überein. Damit ist die Fragestellung auf M zurückgeführt. Wir betrachten den Fall orientierbarer Flächen. Der Fall beliebiger Flächen ergibt sich aus dem folgenden Beweis durch Vereinfachungen.

Eine Fläche auf M heiße jetzt charakteristisch, wenn sie orientierbar ist, wenn ihr Rand eine Kurve auf $\mathfrak{T}$ ist, die m in genau einem Punkte schneidet, und wenn sie unter allen orientierbaren Flächen mit solchem Rand von maximaler Charakteristik ist. Dabei bleibe offen, ob es auf M orientierbare Flächen mit einem Rand der angegebenen Art gibt. Dies könnte mit homologischen Hilfsmitteln entschieden werden, wird aber hier nicht benötigt.

Es sind nur die Aussagen I, II und III von 2.6 (Normalisierung, Reduktion, Signifikanz) zu beweisen.

Nehmen wir an, daß I und II bei hypothetischer Existenz von charakteristischen Flächen bereits bewiesen sind, so ergibt sich die Signifikanz auf einfache Weise. Wenn es überhaupt orientierbare Flächen gibt, deren Rand die gewünschten Eigenschaften hat, so muß es charakteristische Flächen geben

[12]) Wegen der Gestalt von $\mathfrak{V}$ ist ein solcher Kreisring aufweisbar.

und unter diesen wegen I und II Fundamentalflächen. Bis auf unwesentliche Isotopie werden alle Fundamentalflächen durch einen endlichen, repräsentativen Flächenvorrat erfaßt, der aufgewiesen werden kann. Für jede Fläche dieses Vorrats läßt sich nachprüfen, ob sie orientierbar ist und ob ihr Rand die gewünschte Gestalt hat. Findet man solche Flächen in dem Vorrat nicht auf, so gibt es keine charakteristischen Flächen. Findet man welche auf, so lassen sich diejenigen aussondern, die von maximaler Charakteristik sind. Sie sind charakteristisch.

Es bleiben noch I und II zu beweisen.

5.3. Normalisierung. Wir nehmen an, daß es eine charakteristische Fläche $\mathfrak{F}$ gibt. Es ist zu zeigen, daß die einzelnen Schritte in 2.2 ausführbar sind. Dabei kann man sich auf die Schritte 3, 5, 7 und 8 beschränken; alle anderen sind Isotopien bei festem Rand.

Der 3. Schritt fegt $\dot{\mathfrak{F}}$ aus den Flächenstücken von $\mathfrak{T}$ aus, in denen Platten anstoßen. Weil m ein solches Flächenstück höchstens einmal kreuzt, kann dieses Ausfegen so vorgenommen werden, daß sich $\dot{\mathfrak{F}}$ und m stets nur in einem Punkte schneiden, was übrigens nur für die Endlage wesentlich ist.

Der 5. Schritt beseitigt Schleifen von $\dot{\mathfrak{F}}$, die auf Randseen liegen oder einen Randsee mit sich selbst verbinden. m meidet die Randseen und hat mit $\dot{\mathfrak{F}}$ nur einen Schnittpunkt gemein. Hieraus folgt, daß m keine Schleife von $\dot{\mathfrak{F}}$ trifft und daß jede solche Schleife durch Deformation über ein zu m punktfremdes Flächenstück beseitigt werden kann.

Der 7. Schritt ist ausführbar nach dem in 2.2 angegebenen Beispiel 3. Daß der Rand hier nicht fest vorgegeben ist, spielt keine Rolle, da $\dot{\mathfrak{F}}$ bei diesem Schritt ungeändert bleibt.

Bei dem 8. Schritt wird $\mathfrak{F}$ längs eines Querschnittes aufgeschnitten, und an die beiden entstehenden Randbögen werden Elementarflächenstücke angeheftet. Dabei geht $\mathfrak{F}$ in ein Flächensystem $\mathfrak{F}^+$ über, das aus ein oder zwei Flächen besteht. $\mathfrak{F}^+$ ist orientierbar, und es ist $\chi(\mathfrak{F}^+)=\chi(\mathfrak{F})+1$. Der Rand $\dot{\mathfrak{F}}^+$ von $\mathfrak{F}^+$ hat ebenso wie $\dot{\mathfrak{F}}$ nur einen Schnittpunkt mit m gemein, weil $\dot{\mathfrak{F}}$ nur in Grenznähe eines Randsees geändert wird und m die Randseen meidet. Wäre $\dot{\mathfrak{F}}^+$ zusammenhängend, so wäre $\mathfrak{F}^+$ eine einzige Fläche, die geeigneten Rand, aber zu große Charakteristik hätte [13]).

$\dot{\mathfrak{F}}^+$ muß also aus zwei Kurven t_1 und t_2 auf $\mathfrak{T}$ bestehen, wovon eine, etwa t_1, den Meridian m in einem Punkte schneidet, während die andere, also t_2, weder m noch t_1 trifft. m und t_1 bilden ein Paar konjugierter Rückkehrschnitte auf $\mathfrak{T}$. Daher ist t_2 Rand eines Elementarflächenstückes $\mathfrak{h}$ auf $\mathfrak{T}$, das m und t_1 nicht enthalten kann. $\mathfrak{h}$ werde zu $\mathfrak{F}^+$ hinzugenommen und von $\mathfrak{T}$ ins Innere von M abgehoben. Dabei bleibt die Orientierbarkeit erhalten, die Charakteristik erhöht sich nochmals um 1, und der Rand ist wieder eine Kurve der gewünschten Gestalt. Weil $\mathfrak{F}$ bei solchem Rand von maximaler Charakteristik

[13]) Dieser Schluß bleibt bestehen, wenn $\mathfrak{F}$ nicht-orientierbar ist.

ist, müssen jetzt zwei Flächen vorliegen, wovon eine mindestens die Charakteristik 2 haben muß, also eine 2-Sphäre ist, während die andere eine charakteristische Fläche ist, die nun an die Stelle von $\mathfrak{F}$ tritt (insgesamt wurde von $\mathfrak{F}$ ein Elementarflächenstück abgeschnitten und ein anderes dafür angefügt). Der 8. Schritt ist also ausführbar und damit die gesamte Normalisierung.

Wir merken an, daß wir das eben angegebene Verfahren, eine zusätzliche Randkurve unter Erhöhung der Charakteristik zu beseitigen, weiter unten noch mehrfach benutzen.

5.4. Reduktion. Wir nehmen an, daß es eine charakteristische Normalfläche $\mathfrak{F}$ gibt, die nicht fundamental ist. $(\mathfrak{F}_1, \mathfrak{F}_2)$ sei ein erzeugendes Normalflächenpaar für $\mathfrak{F}$ gemäß 3.4. Es ist möglich, daß $\mathfrak{F}_1$ und $\mathfrak{F}_2$ berandet sind und mehrere Randkurven besitzen. Nur auf einer dieser Randkurven liegt ein Schnittpunkt mit m. Diese Randkurve gehöre zu $\mathfrak{F}_1$. $\mathfrak{F}_1$ kann nichtorientierbar sein. Es werde $\mathfrak{F}$ mit einer Orientierung versehen. Zerschneiden wir $\mathfrak{F}_1$ und $\mathfrak{F}_2$ längs ihrer Schnitte, so sind die entstehenden Teile orientierbar. Sie mögen mit der von $\mathfrak{F}$ herrührenden Orientierung versehen sein. Je nach Gestalt dieser Teile unterscheiden wir drei Fälle.

1. Fall. Unter den Teilen von $\mathfrak{F}_1$ und $\mathfrak{F}_2$ befinden sich Elementarflächenstücke, die von geschlossenen Schnittlinien berandet werden.

Elementarflächenstücke dieser Art nennen wir unzerlegte Elementarflächenstücke von $\mathfrak{F}$.

a) Es kann sein, daß ein unzerlegtes Elementarflächenstück $\mathfrak{e}$ dadurch entsteht, daß eine projektive Ebene längs eines orientierbarmachenden Rückkehrschnittes s aufgeschnitten wird. Dann ist $\mathfrak{F}_2$ eine projektive Ebene und s der einzige Schnitt von $\mathfrak{F}_1$ und $\mathfrak{F}_2$. Wegen 3.4, (6) ist s orientierungsumkehrend auf $\mathfrak{F}_1$ und auf $\mathfrak{F}_2$, und insbesondere zerlegt s weder $\mathfrak{F}_1$ noch $\mathfrak{F}_2$. Wir schalten s irregulär um. Es entsteht eine einzige Fläche $\tilde{\mathfrak{F}}$, die orientiert ist und mit $\mathfrak{F}$ in Rand und Charakteristik übereinstimmt. Gemäß 3.5 läßt sich durch Normalisieren von $\tilde{\mathfrak{F}}$ eine charakteristische Normalfläche erhalten, die einfacher als $\mathfrak{F}$ ist.

b) Die unzerlegten Elementarflächenstücke von $\mathfrak{F}$ seien von nun an von Schnittlinien berandet, die $\mathfrak{F}_1$ oder $\mathfrak{F}_2$ zerlegen. Sei $\mathfrak{e}$ ein solches und s sein Rand. Wir schalten alle Schnitte von $\mathfrak{F}_1$ und $\mathfrak{F}_2$ außer s regulär um und zerschneiden längs s. Dann müssen genau 3 berandete Flächenteile von $\mathfrak{F}$ vorliegen, nämlich $\mathfrak{e}$, ein Teil $\mathfrak{F}'$, der s nur einmal im Rande enthält, und ein Teil $\mathfrak{F}''$, der zwei von s herrührende Randkurven besitzt. Jedenfalls müssen aus s vier Randkurven entstehen, weil s nicht orientierungsumkehrend ist. Diese Randkurven können sich nicht auf 4 Flächenteile verteilen, weil sonst die reguläre Umschaltung von s nicht $\mathfrak{F}$ sondern zwei Flächen ergäbe. Es können auch nicht nur 2 Flächenteile vorliegen, sonst besäße der von $\mathfrak{e}$ verschiedene Teil drei von s herrührende Randkurven, die durch drei Exemplare von $\mathfrak{e}$ unter gegenseitigem Abheben verschlossen werden könnten. Es entstände eine orientierbare Fläche, die denselben Rand wie $\mathfrak{F}$, aber größere Charakteristik hätte, was unmöglich ist.

Liegt $\dot{\mathfrak{F}}$ auf $\mathfrak{F}'$, so bilden $\mathfrak{F}'$ und $\mathfrak{e}$ eine orientierbare Fläche $\tilde{\mathfrak{F}}$, die $\dot{\mathfrak{F}}$ als Rand besitzt und deren Charakteristik nicht kleiner als $\chi(\mathfrak{F})$ sein kann. Denn es ist $\chi(\mathfrak{F}'') \leqq 0$ wegen der beiden Randkurven von $\mathfrak{F}''$, und es ist $\chi(\mathfrak{F}) = \chi(\mathfrak{e}) + \chi(\mathfrak{F}') + \chi(\mathfrak{F}'')$. Also ist $\tilde{\mathfrak{F}}$ charakteristisch. Gemäß 3.5 führt $\tilde{\mathfrak{F}}$ auf eine charakteristische Normalfläche, die einfacher als $\mathfrak{F}$ ist.

c) Es kann weiterhin angenommen werden, daß $\dot{\mathfrak{F}}$ zu $\mathfrak{F}''$ gehört. Fügt man nun vorübergehend je ein Exemplar von $\mathfrak{e}$ in die beiden zu s gehörigen Randkurven von $\mathfrak{F}''$ ein und hebt man diese beiden Exemplare voneinander ab, so entsteht eine orientierbare Fläche mit Rand $\dot{\mathfrak{F}}$. Ihre Charakteristik kann nicht größer als $\chi(\mathfrak{F})$ sein, weil $\mathfrak{F}$ charakteristisch ist. Daraus folgt, daß $\mathfrak{F}'$ ein Elementarflächenstück sein muß. Ist $\mathfrak{F}'$ unzerlegt, so muß $\mathfrak{F}_2$ aus $\mathfrak{e}$ und $\mathfrak{F}'$ bestehen[14]), weil s andernfalls $\mathfrak{F}_1$ und $\mathfrak{F}_2$ zugleich zerlegte, was wegen 3.4, (5) ausgeschlossen ist. Außerdem muß s die einzige Schnittlinie von $\mathfrak{F}_1$ und $\mathfrak{F}_2$ sein. In diesem Falle erhält man durch irreguläres Umschalten von s eine charakteristische Fläche $\tilde{\mathfrak{F}}$, und aus ihr durch Normalisieren gemäß 3.5 eine einfachere Fläche als $\mathfrak{F}$.

d) Es bleibt der Fall zu betrachten, daß sich $\mathfrak{F}'$ aus mehreren Teilen von $\mathfrak{F}_1$ und $\mathfrak{F}_2$ zusammensetzt. Wir nennen dann $\mathfrak{F}'$ das zu $\mathfrak{e}$ gehörige zerlegte Elementarflächenstück. Auf $\mathfrak{F}'$ muß ein unzerlegtes Elementarflächenstück von $\mathfrak{F}$ liegen. Nach der bisherigen Diskussion können wir annehmen, daß es zu jedem unzerlegten Elementarflächenstück $\mathfrak{e}$ von $\mathfrak{F}$ ein zugehöriges zerlegtes $\mathfrak{e}'$ gibt. Da $\mathfrak{F}_1$ und $\mathfrak{F}_2$ durch ihre Schnitte in nur endlich viele Teil zerfallen, gibt es unzerlegte Elementarflächenstücke $\mathfrak{e}_1, \mathfrak{e}_2, \ldots, \mathfrak{e}_r$ derart, daß $\mathfrak{e}_{i+1}$ auf dem zu $\mathfrak{e}_i$ gehörigen zerlegten Elementarflächenstück $\mathfrak{e}_i'$ liegt ($i = 1, 2, \ldots, r$; Indices modulo r). $s_1, s_2, \ldots, s_r$ seien die Schnittlinien von $\mathfrak{F}_1$ und $\mathfrak{F}_2$, die $\mathfrak{e}_1, \mathfrak{e}_2, \ldots, \mathfrak{e}_r$ ausschneiden. Wir schalten alle Schnitte regulär um, lassen aber $s_1, s_2, \ldots, s_r$ als Berührungslinien bestehen. Danach lassen wir die zerlegten Elementarflächenstücke $\mathfrak{e}_1', \mathfrak{e}_2', \ldots, \mathfrak{e}_r'$ fort und fügen an die entstehenden Ränder $\mathfrak{e}_1, \mathfrak{e}_2, \ldots, \mathfrak{e}_r$ an, soweit sie benötigt werden. Es entsteht eine charakteristische Fläche $\tilde{\mathfrak{F}}$, aus der man wegen 3.5 durch Normalisieren eine einfachere Fläche als $\mathfrak{F}$ erhält[15]).

Damit ist der 1. Fall vollständig erledigt.

2. Fall. Der 1. Fall tritt nicht ein. Unter den Teilen von $\mathfrak{F}_1$ und $\mathfrak{F}_2$ befinden sich Elementarflächenstücke, die von $\mathfrak{F}_1$ oder $\mathfrak{F}_2$ durch einen einzigen Querschnitt abgeschnitten werden und auf deren Rand nicht der Schnittpunkt von $\dot{\mathfrak{F}}$ und m liegt.

Unter unzerlegten Elementarflächenstücken von $\mathfrak{F}$ wollen wir jetzt solche Elementarflächenstücke verstehen. Die Diskussion entspricht der des 1. Falles.

a) Elementarflächenstücke, die durch Aufschneiden eines Möbius-Bandes oder Kreisringes längs eines einzigen Querschnittes entstehen, zählen nicht

[14]) $\mathfrak{F}_1$ kommt wegen seines Randes nicht in Frage.

[15]) Tatsächlich ist $\tilde{\mathfrak{F}}$ bereits Normalfläche.

zu den unzerlegten. Daher ist kein Analogon zum Abschnitt a) des 1. Falles vorhanden.

b) Es sei e ein unzerlegtes Elementarflächenstück und s die Schnittlinie, die e von $\mathfrak{F}_1$ oder $\mathfrak{F}_2$ abschneidet. Nach Abschnitt a) besitzt e nur einen Randbogen, der von s herrührt. Wir schalten alle von s verschiedenen Schnittlinien von $\mathfrak{F}_1$ und $\mathfrak{F}_2$ regulär um und zerschneiden längs s, womit sich die Gesamtcharakteristik um 2 erhöht. Es müssen nun genau drei Flächenteile vorliegen, nämlich e, ein Teil $\mathfrak{F}'$, der s nur einmal im Rande enthält, und ein Teil $\mathfrak{F}''$, der zwei von s herrührende Randbögen besitzt. Wären 4 Teile vorhanden, so würden durch reguläres Umschalten von s zwei Flächen und nicht $\mathfrak{F}$ entstehen. Wären nur 2 Teile vorhanden, so besäße der von e verschiedene Teil drei von s herrührende Randbögen. An diese könnten drei Exemplare von e angesetzt werden (die gleichzeitig voneinander abgehoben werden), wodurch eine orientierbare Fläche $\mathfrak{F}^+$ mit $\chi(\mathfrak{F}^+)=\chi(\mathfrak{F})+1$ entstände. Der Rand von $\mathfrak{F}^+$ hätte wieder nur einen Schnittpunkt mit m gemein, weil m den Rand von e nicht trifft. Die Annahme, daß $\mathfrak{F}^+$ nur eine Randkurve besitzt, ist widerspruchsvoll wegen $\chi(\mathfrak{F}^+)>\chi(\mathfrak{F})$. Die Annahme, daß $\mathfrak{F}^+$ mehrere Randkurven besitzt, ist ebenfalls widerspruchsvoll. Wie in 5.3 folgte, daß diejenigen Randkurven, die m nicht schneiden, Elementarflächenstücke auf $\mathfrak{T}$ beranden. Man könnte daher der Reihe nach solche Elementarflächenstücke zu $\mathfrak{F}^+$ hinzunehmen und von $\mathfrak{T}$ ins Innere von M abheben, bis nur noch eine Randkurve vorliegt. Die Charakteristik würde sich dabei weiter erhöhen, was den Annahmen über $\mathfrak{F}$ widerspricht. Tatsächlich liegen also die Teile e, $\mathfrak{F}'$ und $\mathfrak{F}''$ vor.

Nehmen wir zunächst an, daß m den Rand von $\mathfrak{F}'$ schneidet. Dann bilden $\mathfrak{F}'$ und e zusammen eine orientierbare Fläche $\tilde{\mathfrak{F}}$, deren Rand von m in einem Punkte geschnitten wird. Wegen $\chi(\mathfrak{F}'')\leqq 1$, $\chi(e)+\chi(\mathfrak{F}')+\chi(\mathfrak{F}'')=\chi(\mathfrak{F})+2$ und $\chi(\tilde{\mathfrak{F}})=\chi(\mathfrak{F}')$ muß $\chi(\tilde{\mathfrak{F}})\geqq\chi(\mathfrak{F})$ sein. $\tilde{\mathfrak{F}}$ besitzt nur eine Randkurve, weil man andernfalls aus $\tilde{\mathfrak{F}}$ wie zuvor eine orientierbare Fläche mit nur einer Randkurve und größerer Charakteristik als $\mathfrak{F}$ erhalten könnte. $\tilde{\mathfrak{F}}$ ist also charakteristisch und führt gemäß 3.5 auf eine einfachere Fläche als $\mathfrak{F}$.

c) Es kann nunmehr angenommen werden, daß m den Rand von $\mathfrak{F}''$ schneidet. Fügt man vorübergehend an die beiden zu s gehörigen Randbögen zwei Exemplare von e an (die man gleichzeitig voneinander abhebt), so muß eine Fläche entstehen, deren Charakteristik nicht größer als $\chi(\mathfrak{F})$ sein kann, weil sich sonst wie zuvor ein Widerspruch zu den Annahmen über $\mathfrak{F}$ ergäbe. Damit folgt, daß $\mathfrak{F}'$ ein Elementarflächenstück ist. Ist $\mathfrak{F}'$ unzerlegt, so muß $\mathfrak{F}_2$ aus $\mathfrak{F}'$ und e bestehen und s die einzige Schnittlinie von $\mathfrak{F}_1$ und $\mathfrak{F}_2$ sein, weil s andernfalls $\mathfrak{F}_1$ und $\mathfrak{F}_2$ zugleich zerlegte, was wegen 3.4, (5) ausgeschlossen ist. Schaltet man s irregulär um, so erhält man eine charakteristische Fläche $\tilde{\mathfrak{F}}$ und aus ihr durch Normalisieren eine einfachere Fläche als $\mathfrak{F}$.

d) Es bleibt der Fall zu betrachten, daß sich $\mathfrak{F}'$ aus mehreren Teilen von $\mathfrak{F}_1$ und $\mathfrak{F}_2$ zusammensetzt. Wir nennen $\mathfrak{F}'$ das zu e gehörige zerlegte Elementarflächenstück. Weil der 1. Fall ausgeschlossen ist, setzen sich die Teile von

$\mathfrak{F}'$ längs Querschnitten von $\mathfrak{F}'$ zusammen, und es gibt auf $\mathfrak{F}'$ ein unzerlegtes Elementarflächenstück. Wegen der Abschnitte b) und c) kann angenommen werden, daß es zu jedem zerlegten Elementarflächenstück ein unzerlegtes gibt. Man kann nun wie im Abschnitt d) des 1. Falles weiterschließen. Man erhält eine Fläche $\tilde{\mathfrak{F}}$ mit $\chi(\tilde{\mathfrak{F}}) = \chi(\mathfrak{F})$. Der Rand von $\tilde{\mathfrak{F}}$ schneidet m in genau einem Punkte, weil m kein zerlegtes Elementarflächenstück trifft. Der Rand von $\tilde{\mathfrak{F}}$ besteht wieder nur aus einer Kurve, weil sich sonst wie zuvor ein Widerspruch zu den Annahmen über $\mathfrak{F}$ ergäbe. $\tilde{\mathfrak{F}}$ ist also charakteristisch und führt auf eine einfachere Fläche als $\mathfrak{F}$.

Damit ist auch der 2. Fall vollständig erledigt.

3. Fall. Es liegen weder der 1. noch der 2. Fall vor.

Wir stellen zunächst fest, daß $\chi(\mathfrak{F}_2) \leqq 0$ und daher $\chi(\mathfrak{F}_1) \geqq \chi(\mathfrak{F})$ ist. Wäre nämlich $\mathfrak{F}_2$ eine 2-Sphäre oder eine projektive Ebene, so müßte der 1. Fall vorliegen; wäre $\mathfrak{F}_2$ ein Elementarflächenstück, so müßte der 1. oder der 2. Fall vorliegen.

Ist nun $\mathfrak{F}_1$ orientierbar, so kann $\mathfrak{F}$ durch die einfachere Fläche $\mathfrak{F}_1$ ersetzt werden. $\mathfrak{F}_1$ kann nämlich nur eine Randkurve besitzen, weil sich sonst wie zuvor ein Widerspruch zu den Annahmen über $\mathfrak{F}$ ergäbe.

Sei jetzt $\mathfrak{F}_1$ nicht-orientierbar. Weil die Teile von $\mathfrak{F}_1$ und $\mathfrak{F}_2$ orientiert sind, gibt es eine Schnittlinie s von $\mathfrak{F}_1$ und $\mathfrak{F}_2$ derart, daß die beiden Ufer[16]) von s auf $\mathfrak{F}_1$ in unverträglicher Weise orientiert sind. Die beiden Ufer von s auf $\mathfrak{F}_2$ sind ebenfalls in unverträglicher Weise orientiert, weil sich bei regulärem Umschalten von s die Orientierung von $\mathfrak{F}$ ergibt. Wir schalten alle Schnitte außer s regulär um und s irregulär. Es entstehen eine oder zwei orientierte Flächen. Entsteht nur eine, so muß diese aus mehrfach erwähnten Gründen charakteristisch sein, und man erhält aus ihr durch Normalisieren eine einfachere Fläche als $\mathfrak{F}$. Es mögen nun zwei Flächen $\tilde{\mathfrak{F}}$ und $\tilde{\tilde{\mathfrak{F}}}$ entstehen. m schneide den Rand von $\tilde{\mathfrak{F}}$. Wie oben für $\mathfrak{F}_2$ folgt für $\tilde{\tilde{\mathfrak{F}}}$ nun $\chi(\tilde{\tilde{\mathfrak{F}}}) \leqq 0$. Damit ergibt sich wieder, daß $\tilde{\mathfrak{F}}$ charakteristisch sein muß, so daß man durch Normalisieren eine einfachere Fläche als $\mathfrak{F}$ erhalten kann.

Damit ist die Reduktion für alle Fälle ausgeführt und der Beweis des Satzes in 5.1 beendet.

5.5. Flächen mit vorgegebenem Rand.

Satz. *Es sei M eine berandete, 3-dimensionale, kompakte Mannigfaltigkeit und $\dot{M}$ ihr Rand. C sei eine einfache geschlossene Kurve auf $\dot{M}$, die Unterkomplex einer gegebenen regulären Zellenzerlegung Σ von M ist. Es läßt sich entscheiden, ob C auf M orientierbare Flächen berandet. Ist dies der Fall, so läßt sich eine von C berandete, orientierbare Fläche maximaler Charakteristik aufweisen. Entsprechende Aussagen gelten, wenn orientierbare Flächen durch beliebige Flächen ersetzt werden.*

[16]) Ufer in lokalem Sinne. s kann orientierungsumkehrend, also einufrig sein.

Der Beweis ergibt sich aus dem Vorangehenden durch Vereinfachung. Weil C Unterkomplex von $\dot{M}$ bezüglich Σ ist, sind bei der Normalisierung die den Rand betreffenden Schritte in 5.3 gegenstandslos (vgl. die Anmerkung am Ende von 2.2). Bei der Reduktion ist in 5.4 für $\mathfrak{F}_1$ diejenige Fläche zu nehmen, die von C berandet wird. $\mathfrak{F}_2$ ist dann geschlossen, und der 2. Fall ist gegenstandslos. Die Signifikanz ergibt sich entsprechend 5.2, wobei die Bemerkung in 2.6 bezüglich des vorgegebenen Randes zu beachten ist.

§ 6. Die Primfaktorzerlegung von Verkettungen

6.1. Faktorzerlegung. Es sei V eine orientierte Verkettung in der S^3. Jede Komponente von V ist ein orientierter Polygonzug. Orientierte Knoten sind als spezielle orientierte Verkettungen aufzufassen. Es wird vorausgesetzt, daß V nicht zerfällt. Nach § 3 ist dies entscheidbar, wenn V gegeben ist. $\mathfrak{S}$ sei eine 2-Sphäre, die V in genau zwei Punkten schneidet und sonst nicht trifft (Fig. 17a). Weil die S^3 von $\mathfrak{S}$ zerlegt wird, liegen die beiden Schnitt-

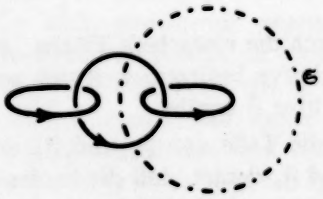

Fig. 17 a

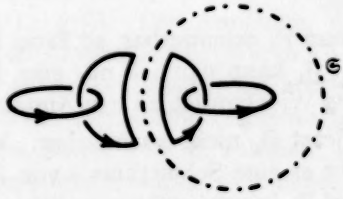

Fig. 17 b

punkte A und B von V mit $\mathfrak{S}$ auf derselben Komponente von V. Sei k diese Komponente. w sei ein einfacher Weg auf $\mathfrak{S}$, der A und B verbindet. Es ist unwesentlich, wie w gewählt wird, da je zwei solche Wege kombinatorisch isotop auf $\mathfrak{S}$ sind. k werde in A und B aufgeschnitten. An jeden der beiden entstehenden Teile von k wird w mit passender Orientierung angefügt. Werden die beiden Exemplare von w nach verschiedenen Seiten von $\mathfrak{S}$ geeignet abgehoben, so entstehen zwei orientierte Verkettungen V_1 und V_2, die durch $\mathfrak{S}$ getrennt werden (Fig. 17b). Aus k ist dabei je eine Komponente k_1 bzw. k_2 von V_1 bzw. V_2 entstanden. Durch V und $\mathfrak{S}$ sind V_1 und V_2 bis auf semilineare Äquivalenz bestimmt.

Wir sagen, daß V und $\mathfrak{S}$ eine Produktdarstellung von V bezüglich k bilden oder daß V durch $\mathfrak{S}$ in die Faktoren V_1 und V_2 zerlegt wird. V heißt das Produkt von V_1 und V_2 bezüglich k_1 und k_2. Durch die Faktoren V_1, V_2 und die Angabe der Komponenten k_1, k_2 ist V bis auf semilineare Äquivalenz eindeutig bestimmt. Das Produkt heißt trivial, wenn V_1 oder V_2 ein (orientierter) Kreis ist. Entsprechend wird ein Faktor als trivial bezeichnet, wenn er ein Kreis ist. Ein Kreis läßt sich nur als triviales Produkt darstellen. Ist V kein Kreis und läßt sich V nur als triviales Produkt darstellen, so heißt V prim. Ist V nicht-triviales Produkt von V_1 und V_2 und ist etwa V_1 nicht prim, so kann V_1 weiter in nicht-trivialer Weise in Faktoren zerlegt werden. V ist dann als Produkt mehrerer nicht-trivialer Faktoren aufzufassen. Diese

Produktbildung ist assoziativ und kommutativ. Kreise fassen wir als leere Produkte von primen Verkettungen auf, prime Verkettungen als Produkte primer Verkettungen mit nur einem Faktor. In Verallgemeinerung des Sachverhaltes bei Knoten wurde von Y. HASHIZUME [2] bewiesen: Jede nichtzerfallende, orientierte Verkettung V ist Produkt von endlich vielen primen Verkettungen, die durch V bis auf semilineare Äquivalenz eindeutig bestimmt sind. Umgekehrt ist V durch seine Primfaktoren und die Angabe, bezüglich welcher Komponentenpaare die Produkte zu bilden sind, bis auf semilineare Äquivalenz eindeutig bestimmt.

6.2. Satz. *In der S^3 sei eine nicht-zerfallende, orientierte Verkettung V gegeben. Die Primfaktorzerlegung von V kann aufgewiesen werden.*

Bezüglich der Vorgabe einer Verkettung in der S^3 gelten die Bemerkungen in 4.1.

Ist eine orientierte Verkettung V in der S^3 gegeben, so läßt sich wegen § 4 entscheiden, ob V zerfällt. Es kann also zunächst nachgeprüft werden, ob die Voraussetzung des Satzes erfüllt ist. Sei dies von nun an der Fall. Wegen § 5 läßt sich feststellen, ob V ein Kreis ist. Ist dies der Fall, so ist nichts weiter aufzuweisen. Wir können dies also im folgenden ausschließen.

Wir reduzieren zunächst die Behauptung des Satzes auf eine einfachere Aussage. Es sei k eine Komponente von V. Wir fragen, ob es eine nichttriviale Produktdarstellung von V bezüglich k gibt. Liegt ein Verfahren vor, diese Frage zu entscheiden und bei positiver Antwort eine solche Produktdarstellung durch Angabe einer 2-Sphäre zu V aufzuweisen, so ergibt sich die Behauptung des Satzes folgendermaßen:

Man führe das Verfahren der Reihe nach für die verschiedenen Komponenten von V aus. Man findet entweder eine nicht-triviale Produktdarstellung von V auf, oder es ergibt sich, daß V prim ist. Ist eine 2-Sphäre $\mathfrak{S}$ aufgewiesen, die V in nur zwei Punkten schneidet und V als nicht-triviales Produkt darstellt, so bewirkt $\mathfrak{S}$ eine Unterteilung der vorhandenen Zellenzerlegung der S^3. Jeder der beiden durch $\mathfrak{S}$ aufgewiesenen Faktoren läßt sich als Unterkomplex der S^3 bezüglich der Unterteilung angeben. Damit ist das Verfahren auf die Faktoren erneut anwendbar, so daß man nach endlich vielen Wiederholungen zu den Primfaktoren von V gelangt.

Die Anzahl der erforderlichen Wiederholungen läßt sich übrigens von vornherein abschätzen, da man eine obere Schranke für die Anzahl der Primfaktoren angeben kann. Bezeichnet $m(V)$ die Anzahl der Komponenten von V und $g(V)$ das Geschlecht von V, so ist $m(V)+g(V)-1$ eine Invariante von V, die additiv bezüglich des Produktes ist. Sie ist außerdem nicht-negativ und nur dann Null, wenn V ein Kreis ist. $m(V)+g(V)-1$ ist also eine obere Schranke für die Anzahl der Primfaktoren von V. Liegt von V eine normierte Projektion vor, so läßt sich das Geschlecht nach oben abschätzen, indem man nach dem Verfahren von SEIFERT [7] orientierbare Flächen in V einspannt.

6.3. Beweisansatz. Sei weiterhin V eine orientierte Verkettung, die Unterkomplex einer gegebenen Zellenzerlegung der S^3 ist. V zerfalle nicht und sei kein Kreis. k sei eine Komponente von V. Nach dem Vorangehenden muß gezeigt werden: Ist V nicht-triviales Produkt bezüglich k, so läßt sich eine 2-Sphäre aufweisen, die V in zwei nicht-triviale Faktoren zerlegt.

Wir bohren nun V aus der S^3 aus, wodurch eine berandete Mannigfaltigkeit M entsteht. Vermöge der gegebenen Zellenzerlegung der S^3 und der Ränder der ausgebohrten Vollringe erhält man eine reguläre Zellenzerlegung Σ von M. $\mathfrak{V}$ sei unter den ausgebohrten Vollringen derjenige, der k zur Seele hat. $\mathfrak{T}$ sei der Rand von $\mathfrak{V}$. Beim Ausbohren von $\mathfrak{V}$ lassen sich zwei disjunkte Meridianflächen $\mathfrak{m}_1$ und $\mathfrak{m}_2$ von $\mathfrak{V}$ so aufweisen, daß sie k in nur je einem Punkte treffen und daß ihre Ränder Unterkomplexe von $\mathfrak{T}$ bezüglich Σ sind. m_1 und m_2 seien die Ränder von $\mathfrak{m}_1$ und $\mathfrak{m}_2$.

Ist $\mathfrak{S}$ irgendeine 2-Sphäre, die V in zwei Punkten von k schneidet und sonst nicht trifft, so läßt sich $\mathfrak{S}$ durch eine orientierungserhaltende semilineare Abbildung der S^3, die V auf sich abbildet, in eine 2-Sphäre überführen, die aus $\mathfrak{m}_1$, $\mathfrak{m}_2$ und einem Kreisring auf M besteht[17]). Bis auf semilineare Äquivalenz erhält man also alle Faktorzerlegungen von V bezüglich k durch Kreisringe, die auf M liegen und von m_1 und m_2 berandet werden, indem man diese Kreisringe durch $\mathfrak{m}_1$ und $\mathfrak{m}_2$ zu 2-Sphären ergänzt. Es genügt daher, solche Kreisringe auf M zu betrachten. Von ihnen sagen wir weiterhin, daß sie V als Produkt darstellen. Eine solche Darstellung heißt nicht-trivial, wenn die entsprechende 2-Sphäre eine nicht-triviale Produktdarstellung von V ergibt.

Für M charakteristische Flächen seien nun Kreisringe auf M, die von m_1 und m_2 berandet werden und die V als nicht-triviales Produkt darstellen.

Durch das Vorangehende und durch 2.6 ist der Beweis des Satzes zurückgeführt auf den Beweis der Aussagen I, II und III von 2.6 (Normalisierung, Reduktion, Signifikanz).

6.4. Normalisierung. Wir nehmen an, daß es eine charakteristische Fläche $\mathfrak{F}$ gibt. Ihr Rand $\dot{\mathfrak{F}}$ ist Unterkomplex von $\mathfrak{T}$ bezüglich Σ. Daher sind die den Rand betreffenden Schritte in 2.2 gegenstandslos. Es muß nur nachgeprüft werden, daß der 7. Schritt ausführbar ist, denn alle übrigen sind Isotopien bei festem Rand.

Sei dazu $\mathfrak{g}$ ein Elementarflächenstück, dessen Rand s auf $\mathfrak{F}$ liegt und den Durchschnitt von $\mathfrak{g}$ und $\mathfrak{F}$ ausmacht. Schneidet man $\mathfrak{F}$ längs s auf und fügt man an die beiden entstehenden Randkurven je ein Exemplar von $\mathfrak{g}$ an (bei gleichzeitigem Abheben voneinander), so entsteht ein Flächensystem mit der Charakteristik 2 und den beiden Randkurven m_1 und m_2. Es ist ausgeschlossen, daß zwei Elementarflächenstücke entstehen, weil sonst jedes zusammen mit einer der beiden Meridianflächen $\mathfrak{m}_1$, $\mathfrak{m}_2$ eine 2-Sphäre bildete, die nur einen Schnittpunkt mit k besäße. Es liegen also eine 2-Sphäre $\mathfrak{R}$ und ein Kreisring $\tilde{\mathfrak{F}}$ vor. Lassen wir die beiden Exemplare von $\mathfrak{g}$ wieder zusammenfallen, so

[17]) Vgl. [5], § 3.

besteht $\mathfrak{R}$ aus $\mathfrak{g}$ und einem Elementarflächenstück $\mathfrak{f}$ auf $\mathfrak{F}$. $\mathfrak{R}$ berandet in der S^3 eine Kugel $\mathfrak{M}$, die zu V disjunkt ist, weil V nicht zerfällt. Der Durchschnitt von $\mathfrak{M}$ und $\mathfrak{F}$ besteht aus $\mathfrak{f}$. Über $\mathfrak{M}$ hinweg läßt sich $\mathfrak{f}$ in $\mathfrak{g}$ deformieren[18]), wodurch $\mathfrak{F}$ in $\tilde{\mathfrak{F}}$ übergeht. Bis auf semilineare Äquivalenz stellen somit $\mathfrak{F}$ und $\tilde{\mathfrak{F}}$ dieselbe Faktorzerlegung von V her, so daß wir $\mathfrak{F}$ durch $\tilde{\mathfrak{F}}$ ersetzen können. Dies ist gerade der Inhalt des 7. Schrittes. Die Normalisierung ist also ausführbar.

6.5. Reduktion. Wir nehmen an, daß es eine charakteristische Normalfläche $\mathfrak{F}$ gibt, die nicht fundamental ist. $(\mathfrak{F}_1, \mathfrak{F}_2)$ sei ein erzeugendes Normalflächenpaar für $\mathfrak{F}$. Nehmen wir $\mathfrak{m}_1$ und $\mathfrak{m}_2$ zu $\mathfrak{F}_1$ und $\mathfrak{F}_2$ hinzu, so entstehen zwei geschlossene Flächen in der S^3. Hieraus folgt, daß $\mathfrak{F}_1$ und $\mathfrak{F}_2$ orientierbar sind. Außerdem muß eine der beiden Flächen $\mathfrak{F}_1$ und $\mathfrak{F}_2$ geschlossen sein, etwa $\mathfrak{F}_2$, während die andere, also $\mathfrak{F}_1$, von m_1 und m_2 berandet wird. Entweder ist nun $\mathfrak{F}_1$ ein Torus mit zwei Löchern und $\mathfrak{F}_2$ eine 2-Sphäre, oder es ist $\mathfrak{F}_1$ ein Kreisring und $\mathfrak{F}_2$ ein Torus.

1. Fall. $\mathfrak{F}_2$ ist eine 2-Sphäre.

Wir zerschneiden $\mathfrak{F}_1$ längs sämtlicher Schnittlinien mit $\mathfrak{F}_2$. Unter den entstehenden Teilen von $\mathfrak{F}_1$ sei $\mathfrak{F}_1'$ derjenige, der m_1 im Rande enthält. Da $\mathfrak{F}_2$ von jeder Schnittlinie zerlegt wird, ist kein Teil von $\mathfrak{F}_1$ ein Elementarflächenstück (3.4, (5)). Hieraus folgt $\chi(\mathfrak{F}_1') \geqq -2$ wegen $\chi(\mathfrak{F}_1) = -2$.

Diejenigen Randkurven von $\mathfrak{F}_1'$, die auf $\mathfrak{F}_2$ liegen, kann man der Reihe nach durch Elementarflächenstücke auf $\mathfrak{F}_2$ schließen, die man noch von $\mathfrak{F}_2$ abhebt. Nimmt man noch $\mathfrak{m}_1$ hinzu, so kann keine geschlossene Fläche entstehen, weil nur ein Schnittpunkt mit k vorhanden ist. Es folgt, daß auch m_2 zum Rande von $\mathfrak{F}_1'$ gehört.

Wegen $\chi(\mathfrak{F}_1') \geqq -2$ können nun höchstens zwei Randkurven von $\mathfrak{F}_1'$ auf $\mathfrak{F}_2$ liegen. Es müssen auch mindestens zwei sein, weil $\mathfrak{F}_1$ von keiner Schnittlinie allein zerlegt wird. Die beiden Randkurven von $\mathfrak{F}_1'$ auf $\mathfrak{F}_2$ gehören außerdem zu verschiedenen Schnittlinien von $\mathfrak{F}_1$ und $\mathfrak{F}_2$, weil andernfalls beim Aufschneiden von $\mathfrak{F}_1$ nur $\mathfrak{F}_1'$ entstanden wäre, während doch $\mathfrak{F}_2$ die S^3 zerlegt.

Es ergibt sich, daß $\mathfrak{F}_1'$ durch zwei disjunkte Elementarflächenstücke auf $\mathfrak{F}_2$ zu einem Kreisring $\tilde{\mathfrak{F}}$ ergänzt werden kann. Nimmt man $\mathfrak{m}_1$ und $\mathfrak{m}_2$ zu $\tilde{\mathfrak{F}}$ und zu $\mathfrak{F}$ hinzu, so entstehen zwei 2-Sphären $\tilde{\mathfrak{S}}$ und $\mathfrak{S}$. Im Durchschnitt von $\tilde{\mathfrak{S}}$ und $\mathfrak{S}$ können $\mathfrak{m}_1$ und $\mathfrak{m}_2$ durch einen Weg auf $\mathfrak{F}_1'$ verbunden werden. Durch geringfügige Deformation von $\mathfrak{S}$ kann daher erreicht werden, daß der Durchschnitt von $\tilde{\mathfrak{S}}$ und $\mathfrak{S}$ aus einem $\mathfrak{m}_1$ und $\mathfrak{m}_2$ umfassenden Elementarflächenstück $\mathfrak{n}$ und möglicherweise noch aus einfachen Schnittlinien besteht, die paarweise und zu $\mathfrak{n}$ disjunkt sind. Sind Schnittlinien vorhanden, so gibt es eine auf $\tilde{\mathfrak{S}}$ innerste, die auf $\tilde{\mathfrak{S}}$ ein Elementarflächenstück $\tilde{\mathfrak{f}}$ berandet, das $\mathfrak{n}$ und weitere Schnittlinien nicht enthält. Der Rand von $\tilde{\mathfrak{f}}$ schneidet aus $\mathfrak{S}$

[18]) Vgl. etwa [6], S. 138.

ein Elementarflächenstück $\mathfrak{f}$ aus, das $\mathfrak{n}$ nicht enthält. $\tilde{\mathfrak{f}}$ und $\mathfrak{f}$ bilden zusammen eine 2-Sphäre, die in der S^3 eine zu V disjunkte Kugel $\mathfrak{M}$ berandet, weil V nicht zerfällt. Man kann nun $\mathfrak{S}$ dadurch deformieren, daß man $\mathfrak{f}$ über $\mathfrak{M}$ hinweg in $\tilde{\mathfrak{f}}$ deformiert und anschließend so von $\tilde{\mathfrak{f}}$ abhebt, daß mindestens eine Schnittlinie von $\mathfrak{S}$ und $\tilde{\mathfrak{S}}$ verschwunden ist. Auf diese Weise lassen sich alle Schnittlinien von $\mathfrak{S}$ und $\tilde{\mathfrak{S}}$ der Reihe nach beseitigen, so daß dann der Durchschnitt von $\mathfrak{S}$ und $\tilde{\mathfrak{S}}$ nur noch aus $\mathfrak{n}$ besteht. Danach läßt sich durch eine entsprechende Deformation erreichen, daß $\mathfrak{S}$ in $\tilde{\mathfrak{S}}$ übergeht. Es stellen also $\mathfrak{S}$ und $\tilde{\mathfrak{S}}$ und damit $\mathfrak{F}$ und $\tilde{\mathfrak{F}}$ dieselbe Faktorzerlegung von V dar (bis auf semilineare Äquivalenz). Daher kann $\mathfrak{F}$ durch $\tilde{\mathfrak{F}}$ ersetzt werden. Gemäß 3.5 erhält man aus $\tilde{\mathfrak{F}}$ durch Normalisieren eine einfachere Fläche als $\mathfrak{F}$.

2. Fall. $\mathfrak{F}_2$ ist ein Torus.

$\mathfrak{F}_1$ ist ein Kreisring, der zusammen mit $\mathfrak{m}_1$ und $\mathfrak{m}_2$ eine 2-Sphäre $\mathfrak{S}_1$ bildet. Jede Schnittlinie von $\mathfrak{F}_1$ und $\mathfrak{F}_2$ zerlegt $\mathfrak{S}_1$ und damit auch $\mathfrak{F}_1$. Wegen 3.4, (5) wird $\mathfrak{F}_2$ von keiner Schnittlinie zerlegt. $\mathfrak{F}_2$ bildet daher Kreisringe, die abwechselnd auf den beiden von $\mathfrak{S}_1$ in der S^3 berandeten Kugeln liegen. Unter den Schnittlinien von $\mathfrak{F}_1$ und $\mathfrak{F}_2$ gibt es innerste auf $\mathfrak{S}_1$, d.h. solche, die auf $\mathfrak{S}_1$ jeweils ein von weiteren Schnitten freies Elementarflächenstück beranden. Ist unter diesen Elementarflächenstücken eines, das keinen Schnittpunkt mit k enthält, so kann es keines geben, das mit k genau einen Schnittpunkt besitzt, weil sich sonst ein Kreisring auf $\mathfrak{F}_2$ und zwei Elementarflächenstücke auf $\mathfrak{S}_1$ zu einer 2-Sphäre zusammensetzten, die nur einen Schnittpunkt mit k gemein hätte. Damit folgt für die Schnitte von $\mathfrak{F}_1$ und $\mathfrak{F}_2$, daß entweder jeder aus $\mathfrak{F}_1$ ein Elementarflächenstück ausschneidet oder daß jeder $\mathfrak{F}_1$ in zwei Kreisringe zerlegt.

Schneidet jede Schnittlinie aus $\mathfrak{F}_1$ ein Elementarflächenstück aus, so kann man $\mathfrak{m}_1$ und $\mathfrak{m}_2$ durch einen Weg auf $\mathfrak{F}_1$ verbinden, der keine Schnittlinie trifft. Nimmt man nun $\mathfrak{m}_1$ und $\mathfrak{m}_2$ auch zu $\mathfrak{F}$ hinzu, so erhält man eine 2-Sphäre $\mathfrak{S}$, in deren Durchschnitt mit $\mathfrak{S}_1$ ein $\mathfrak{m}_1$ und $\mathfrak{m}_2$ umfassendes Elementarflächenstück $\mathfrak{n}$ liegt. Wie im 1. Falle ergibt sich, daß $\mathfrak{S}$ in $\mathfrak{S}_1$ deformiert werden kann. Es folgt, daß $\mathfrak{F}_1$ charakteristisch ist und an die Stelle von $\mathfrak{F}$ treten kann.

Jede Schnittlinie von $\mathfrak{F}_1$ und $\mathfrak{F}_2$ zerlege nun $\mathfrak{F}_1$ in zwei Kreisringe. Ist $\mathfrak{F}_1$ charakteristisch, so kann $\mathfrak{F}$ durch $\mathfrak{F}_1$ ersetzt werden. Sei dies jetzt nicht der Fall. Dann berandet $\mathfrak{S}_1$ in der S^3 eine Kugel $\mathfrak{M}$, so daß der Durchschnitt von V und $\mathfrak{M}$ eine unverknotete Sehne von $\mathfrak{M}$ ist. $\mathfrak{F}_2$ bildet auf $\mathfrak{M}$ Kreisringe. Sei $\mathfrak{r}_2$ ein solcher Kreisring, s_1 und s_2 seien seine beiden Ränder. s_1 und s_2 schneiden aus $\mathfrak{S}_1$ einen Kreisring $\mathfrak{r}_1$ aus, der auf $\mathfrak{F}_1$ liegt. Weil $\mathfrak{r}_2$ die Sehne von $\mathfrak{M}$ nicht trifft und weil diese Sehne in $\mathfrak{M}$ unverknotet ist, beranden $\mathfrak{r}_1$ und $\mathfrak{r}_2$ zusammen einen unverknoteten Vollring $\mathfrak{W}$, der auf $\mathfrak{M}$ liegt und s_1 und s_2 als Breitenkreise besitzt[19]). Liegen auf $\mathfrak{W}$ noch weitere Kreisringe von $\mathfrak{F}_2$, so schneiden diese von $\mathfrak{W}$ Vollringe gleicher Art ab. Es kann daher angenommen werden, daß $\mathfrak{r}_2$ so gewählt ist, daß $\mathfrak{r}_2$ den Durchschnitt von $\mathfrak{W}$

[19]) [6], § 8.

und $\mathfrak{F}_2$ ausmacht. Wenn wir die Schnittlinien s_1 und s_2 regulär umschalten, so wird $\mathfrak{r}_1$ nicht mit beiden Rändern an die an $\mathfrak{r}_2$ anstoßenden Teile von $\mathfrak{F}_2$ angeheftet, weil sonst im Widerspruch zu 3.4, (4) noch immer zwei Flächen vorhanden wären (Fig. 18a, b zeigt den Sachverhalt im Schnitt). Es können auch nicht $\mathfrak{r}_1$ und $\mathfrak{r}_2$ mit beiden Rändern zusammengeheftet werden. Bei regulärem Umschalten von s_1 und s_2 werden vielmehr $\mathfrak{r}_1$ und $\mathfrak{r}_2$ längs einer Schnittlinie, etwa längs s_1, zusammengeheftet, dagegen nicht längs der anderen s_2 (Fig. 18c). Wir schalten nun alle Schnitte von $\mathfrak{F}_1$ und $\mathfrak{F}_2$ außer s_2 regulär

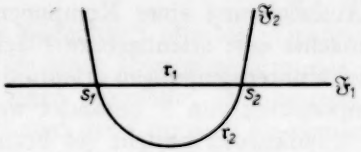

Fig. 18a

Fig. 18b

Fig. 18c

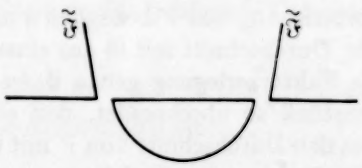

Fig. 18d

um und s_2 irregulär. Dabei entsteht ein aus $\mathfrak{r}_1$ und $\mathfrak{r}_2$ bestehender Torus und ein Kreisring $\tilde{\mathfrak{F}}$ (Fig. 18d). Weil $\mathfrak{W}$ ein Vollring ist, für den s_2 ein Breitenkreis ist, kann $\mathfrak{F}$ über $\mathfrak{W}$ hinweg in $\tilde{\mathfrak{F}}$ deformiert werden[20]) (vgl. Fig. 18c und d). Daher ist $\tilde{\mathfrak{F}}$ charakteristisch, und durch Normalisieren gemäß 3.5 entsteht eine einfachere Fläche als $\mathfrak{F}$.

Die Reduktion ist also in allen Fällen ausführbar.

6.6. Signifikanz. Die Signifikanz ergibt sich hier aus einem bereits vorhandenen Entscheidungsverfahren. Für einen repräsentativen Vorrat von Fundamentalflächen kann zunächst festgestellt werden, welche Flächen von m_1 und m_2 berandete Kreisringe sind[21]). Ist $\mathfrak{r}$ ein solcher Kreisring, so bildet er zusammen mit $\mathfrak{m}_1$ und $\mathfrak{m}_2$ eine 2-Sphäre $\mathfrak{S}$, welche die S^3 in zwei Kugeln zerlegt. Es kann abgelesen werden, ob der Durchschnitt von V mit einer solchen Kugel nur aus einer Sehne besteht. Ist dies für beide Kugeln nicht der Fall, so ist $\mathfrak{r}$ charakteristisch. Besteht der Durchschnitt von V mit einer von $\mathfrak{S}$ berandeten Kugel nur aus einer Sehne, so kann diese durch einen Weg auf $\mathfrak{S}$ zu einem Knoten ergänzt werden. Dies kann so geschehen, daß der Knoten Unterkomplex der S^3 bezüglich einer Zellenzerlegung ist, die durch Unterteilung von Σ' vermöge $\mathfrak{S}$ entsteht. Wegen 5.1 kann dann entschieden werden, ob dieser Knoten ein Kreis ist. Besteht V nur aus einem einzigen Knoten, so ist diese Entscheidung für beide durch $\mathfrak{S}$ aufgewiesenen Faktoren

[20]) [6], § 8.

[21]) Man beachte die Bemerkung über den Rand in 2.6.

zu treffen. Jedenfalls ergibt sich auf endliche Weise, ob $\mathfrak{r}$ charakteristisch ist oder nicht. Damit ist der Satz in 6.2 bewiesen.

6.7. Anmerkung. Das Verfahren läßt sich so anlegen, daß die Frage nach nicht-trivialen Produktdarstellungen von V für alle Komponenten zugleich beantwortet wird. Man hat nur für jeden aus der S^3 ausgebohrten Vollring zwei Meridianflächen anzugeben. Wir haben hierauf verzichtet, um zusätzliche Bezeichnungen zu vermeiden.

Der Satz kann auch noch auf eine andere Weise bewiesen werden, bei der jeweils in Evidenz gesetzt ist, ob eine triviale Faktorzerlegung von V vorliegt oder nicht. Dabei scheint die Auszeichnung einer Komponente k von V wesentlich zu sein. Man kann zunächst eine orientierbare Fläche $\mathfrak{G}$ bestimmen, die von maximaler Charakteristik unter denjenigen orientierbaren Flächen ist, die von k und anderen Komponenten von V berandet werden und deren Durchschnitt mit V nur aus Randkurven besteht (es brauchen nicht alle Komponenten von V erfaßt zu werden). Eine 2-Sphäre $\mathfrak{S}$, die eine Faktorzerlegung von V bezüglich k angibt, kann immer so deformiert werden, daß ihr Durchschnitt mit $\mathfrak{G}$ aus einem einzigen Querschnitt besteht. Danach ist die Faktorzerlegung genau dann trivial, wenn $\mathfrak{S}$ von $\mathfrak{G}$ ein Elementarflächenstück so abschneidet, daß ein Randbogen dieses Elementarflächenstückes den Durchschnitt von V mit einer von $\mathfrak{S}$ berandeten Kugel ausmacht. Um ein Bestimmungsverfahren zu erhalten, schneidet man die S^3 längs $\mathfrak{G}$ auf und bohrt die durch $\mathfrak{G}$ nicht erfaßten Komponenten aus der S^3 aus. Die charakteristischen Flächen sind dann spezielle Elementarflächenstücke.

Literatur

[1] HAKEN, W.: Ein Isotopiekriterium für den Kreisknoten. Erscheint in Acta Math.

[2] HASHIZUME, Y.: On the uniqueness of the decomposition of a link. Osaka Math. J. **10**, 283–300 (1958).

[3] PAPAKYRIAKOPOULOS, C. D.: On the ends of the fundamental groups of 3-manifolds with boundary. Comm. Math. Helv. **32**, 85–92 (1957).

[4] PAPAKYRIAKOPOULOS, C. D.: On DEHN's lemma and the asphericity of knots. Ann. Math. **66**, 1–26 (1957).

[5] SCHUBERT, H.: Die eindeutige Zerlegbarkeit eines Knotens in Primknoten. Akad. Wiss. Heidelberg, math.-nat. Kl., 3. Abh. **1949**.

[6] SCHUBERT, H.: Knoten und Vollringe. Acta math., Stockh. **90**, 131–286 (1953).

[7] SEIFERT, H.: Über das Geschlecht von Knoten. Math. Ann. **110**, 571–592 (1935).

Kiel, Mathematisches Seminar der Universität

(Eingegangen am 10. September 1960)

LOONSTRA, F.
Math. Zeitschr. 76, 149—154 (1961)

Darstellungen von Gruppen als homomorphes Bild

Von

F. LOONSTRA

§ 1. Einleitung

Die Möglichkeit, eine feste Gruppe B als homomorphes Bild einer Gruppe G mit Hilfe eines Homomorphismus α zu beschreiben, liegt den folgenden Betrachtungen zugrunde. Wir nennen $G(\alpha)$ eine *Projektionsdarstellung* (kurz: Darstellung) der Gruppe B als homomorphes Bild, wenn es eine homomorphe Abbildung α von G auf B gibt. Wir untersuchen zuerst die Gesamtheit aller möglichen Darstellungen $G(\alpha)$ von B, und dazu werden[1]) folgende Definitionen eingeführt:

1. Es gibt eine *Ordnungsrelation:*

$G(\alpha) \leqq G'(\alpha')$, wenn es eine homomorphe Abbildung η gibt

$$\eta\colon \quad G' \to G, \text{ so daß } \alpha' = \eta\,\alpha.$$

Die Ordnungsrelation ist transitiv.

2. Zwei Darstellungen $G(\alpha)$ und $G'(\alpha')$ sind *verwandt*, wenn zugleich

$$G(\alpha) \leqq G'(\alpha') \quad \text{und} \quad G'(\alpha') \leqq G(\alpha),$$

d. h. wenn es *verbindende Homomorphismen* η, η' gibt

$$\eta\colon \quad G' \to G \quad \text{mit} \quad \alpha' = \eta\,\alpha,$$
$$\eta'\colon \quad G \to G' \quad \text{mit} \quad \alpha \overset{\circ}{=} \eta'\,\alpha'.$$

Die Verwandtschaft von Darstellungen ist eine Äquivalenzrelation und bewirkt also eine Einteilung der Gesamtheit aller Darstellungen $G(\alpha)$ von B in Klassen $\{G(\alpha)\}$ von untereinander äquivalenten Darstellungen. Die Gesamtheit der Klassen verwandter Darstellungen ist im allgemeinen keine Menge! Sei $G(\alpha)$ eine Darstellung von B, H irgendeine Gruppe, so ist z. B. das direkte Produkt $(G \times H)$ (β) mit $(g, h)\,\beta = g\,\alpha$ eine mit $G(\alpha)$ verwandte Darstellung von B.

3. Sind $G_1(\alpha_1)$ und $G_2(\alpha_2)$ zwei Darstellungen von B, so ist das *Produkt* dieser Darstellungen das *vollständige B-subdirekte Produkt*[2])

$$G(\alpha) = G_1(\alpha_1) \,\tilde{\times}\, G_2(\alpha_2);$$

---

[1]) LOONSTRA, F.; Klassifikation der Darstellungen einer Gruppe. Acta Math. Acad. sc. Hungaricae **11**, 223—229, (1960).

[2]) Herr WIELANDT hat vorgeschlagen, die Bezeichnung „meromorphes subdirektes Produkt" zu ändern und zu sprechen von einem vollständigen B-subdirekten Produkt.

$G(\alpha)$ ist die Gruppe derjenigen Paare $g=(g_1, g_2)$, $g_i \in G_i$ mit $g_1 \alpha_1 = g_2 \alpha_2$, $(g_1, g_2)\alpha = g_1 \alpha_1$. Allgemein: Sei M eine bestimmte Menge von Darstellungen $G_i(\alpha_i)$ von B, so ist das vollständige B-subdirekte Produkt

$$\tilde{\times}_i G_i(\alpha_i) = G(\alpha)$$

die Untergruppe G des vollständigen direkten Produktes $\{\times_i G_i\}^*$ derjenigen Systeme $g=(\ldots, g_i, \ldots, g_j, \ldots)$ mit $g_i \alpha_i = g_j \alpha_j$, wenn $i \neq j$ und $g\alpha = g_i \alpha_i$.

Man beweist leicht, daß die Klasse des Produktes nur durch die Klassen der Faktoren bestimmt wird. Es sei nun erwähnt, daß die Gesamtheit der Klassen von untereinander verwandten Darstellungen eine kleinste Klasse enthält; das ist die Klasse der *identischen* Darstellung $B(\varepsilon)$ von B, wenn mit ε die identische Abbildung von B auf sich bezeichnet wird.

Hinsichtlich der Ordnung der Klassen $\{G(\alpha)\}$ können wir folgendes sagen:

Wenn M eine bestimmte Menge von Darstellungen $G_i(\alpha_i)$ von B ist, so gilt

$$G_i(\alpha_i) \leqq G(\alpha) = \tilde{\times}_i G_i(\alpha_i),$$

und überdies: wenn für eine Darstellung $H(\beta)$ von B gilt, daß

$$G_i(\alpha_i) \leqq H(\beta) \quad \text{für jedes } G_i(\alpha_i) \in M,$$

so gilt auch

$$G(\alpha) = \tilde{\times}_i G_i(\alpha_i) \leqq H(\beta);$$

es hat also das vollständige B-subdirekte Produkt $\tilde{\times}_i G_i(\alpha_i)$ die Eigenschaften der kleinsten oberen Schranke der Darstellungen von M. Man beweist leicht, daß das freie Produkt

$$F(\beta) = *_i G_i(\alpha) \tag{1}$$

mit

$$f\beta = (g_{i_1} \cdot g_{i_2} \cdot \cdots \cdot g_{i_n})\beta = (g_{i_1}\alpha_{i_1}) \cdot \cdots \cdot (g_{i_n}\alpha_{i_n})$$

alle Eigenschaften der größten unteren Schranke besitzt: es ist $F(\beta) \leqq G_i(\alpha_i)$ für alle $G_i(\alpha_i) \in M$ und aus $H(\beta') \leqq G_i(\alpha_i)$ für alle $G_i(\alpha_i)$ folgt $H(\beta') \leqq F(\beta)$. Wir bezeichnen die Gesamtheit aller Klassen von untereinander verwandten Darstellungen $G(\alpha)$ von B mit $V(B)$ und wir sprechen von dem Pseudoverband $V(B)$.

§ 2. Das Problem der Zerlegung

Es sei $B = \times_i C_i$, d.h. B sei das (diskrete) direkte Produkt der Gruppen C_i. Zu jeder Komponente C_i gehört (ebenso wie zu B) ein Pseudoverband $V(C_i)$ der Klassen von Darstellungen von C_i, und wir fragen nach dem Zusammenhang zwischen den Pseudoverbänden $V(B)$ und $V(C_i)$ von B und den direkten Faktoren C_i.

Jede Darstellung $G(\alpha)$ von B induziert auf eine natürliche Weise eine Darstellung $H_i(\beta_i)$ eines direkten Faktors C_i von B; es ist H_i das vollständige

Original von C_i in G (hinsichtlich der Abbildung α) und

$$h_i\beta_i = h_i\alpha = c_i \in C_i;$$

$H_i(\beta_i)$ enthält den Kern K_α der Abbildung α von G auf B und

$$H_i \cap H_j = K_\alpha \quad \text{für} \quad i \neq j.$$

SATZ I. *Wenn für die Darstellungen $G_1(\alpha_1)$ und $G_2(\alpha_2)$ von B gilt, daß $G_1(\alpha_1) \leqq G_2(\alpha_2)$, so gilt auch für die induzierten Darstellungen $H_1(\beta_1)$ und $H_2(\beta_2)$ eines direkten Faktors C von B: $H_1(\beta_1) \leqq H_2(\beta_2)$.*

Es gibt einen Homomorphismus η: $G_2 \to G_1$ mit $\eta\,\alpha_1 = \alpha_2$. Die Abbildung η induziert eine homomorphe Abbildung von H_2 in G_1; aus $h_2\beta_2 = h_2\alpha_2$ ($h_2 \in H_2$) und $h_2\,\eta\,\alpha_1 = h_2\alpha_2 = h_2\beta_2$ folgt, daß $h_2\eta$ dem vollständigen Original von C in G_1 (hinsichtlich α_1) zugehört, d.h. $h_2\eta \in H_1$; η induziert also einen Homomorphismus $\bar{\eta}$: $H_2 \to H_1$ mit $\bar{\eta}\,\beta_1 = \beta_2$. Es ist also $H_1(\beta_1) \leqq H_2(\beta_2)$.

Aus Satz I folgen sofort

SATZ II. *Verwandte Darstellungen $G_1(\alpha_1)$ und $G_2(\alpha_2)$ von B induzieren verwandte Darstellungen $H_1(\beta_1)$ und $H_2(\beta_2)$ eines direkten Faktors C von B.*

SATZ III. *Induzieren die Darstellungen $G(\alpha)$ und $G'(\alpha')$ von B die Darstellungen $H(\beta)$ und $H'(\beta')$ eines direkten Faktors C von $\mathfrak{B}$, so induziert die Klasse $\{G(\alpha) \tilde{\times} G'(\alpha')\}$ die Klasse $\{H(\beta) \tilde{\times} H'(\beta')\}$, und die Klasse $\{G(\alpha) * G'(\alpha')\}$ induziert die Klasse $\{H(\beta) * H'(\beta')\}$.*

Wenn umgekehrt eine Darstellung $H(\beta)$ von C, gegeben ist, so ist auch eine Darstellung $G(\alpha)$ von B zu bestimmen, welche genau $H(\beta)$ als Darstellung von C_i induziert. Definiert man nämlich

$$G(\alpha) = \left\{H \times \left(\underset{j \neq i}{\times} C_j\right)\right\}(\alpha) \tag{5}$$

und für $g = h \cdot c_{j_1} \cdot c_{j_2} \cdot \ldots \cdot c_{j_n}$

$$g\,\alpha = (h\beta) \cdot c_{j_1} \cdot \ldots \cdot c_{j_n} \quad (c_{j_k} \in C_{j_k}),$$

so ist zuerst $G(\alpha)$ eine Darstellung von B und diese Darstellung $G(\alpha)$ induziert für die Komponente C_i von B genau die Darstellung $H(\beta)$; man hat ja die Elemente $g \in G$ zur Verfügung mit $g\,\alpha \in C_i$, das sind genau diejenigen Elemente

$$g = h \cdot c_{j_1} \cdot c_{j_2} \cdot \ldots \cdot c_{j_n} \qquad (c_{j_k} \in C_{j_k} \neq C_i)$$

mit allen $c_{j_k} = e_{j_k}$, und das sind die Elemente von H. Jede Darstellung $G(\alpha)$ von B induziert also auf eine natürliche Weise Darstellungen $H_i(\beta_i)$ der direkten Faktoren C_i von B und umgekehrt; für jede Darstellung $H_i(\beta_i)$ des direkten Faktors C_i von B ist immer eine Darstellung $G(\alpha)$ von B zu bestimmen, welche $H_i(\beta_i)$ induziert.

SATZ IV. *Sind $H_1(\beta_1)$ und $H_2(\beta_2)$ Darstellungen eines direkten Faktors C von B und ist $H_1(\beta_1) \leqq H_2(\beta_2)$, so gilt auch für die induzierten Darstellungen* (5) *von B: $G_1(\alpha_1) \leqq G_2(\alpha_2)$.*

BEWEIS. Es sind $G_1(\alpha_1)=\{H_1\times(\underset{i}{\times} C_i)\}(\alpha_1)$, $G_2(\alpha_2)=\{H_2\times(\underset{i}{\times} C_i)\}(\alpha_2)$, $C_i \neq C$, die induzierten Darstellungen; für die Elemente

$$g_1 = h_1 c_{i_1} \cdot c_{i_2} \cdot \cdots \cdot c_{i_n}, \qquad g_2 = h_2 \cdot c_{i_1} \cdot c_{i_2} \cdot \cdots \cdot c_{i_n}$$

gilt

$$g_1\alpha_1 = (h_1\beta_1) \cdot c_{i_1} \cdot \cdots \cdot c_{i_n}, \qquad g_2\alpha_2 = (h_2\beta_2) \cdot c_{i_1} \cdot \cdots \cdot c_{i_n}.$$

Weiter gibt es einen Homomorphismus η: $H_2 \to H_1$, $\eta\beta_1=\beta_2$. Wir definieren nun den folgenden Homomorphismus

$$\xi: \quad g_2\xi = (h_2 c_{i_1} \cdot \cdots \cdot c_{i_n})\xi = (h_2\eta) \cdot c_{i_1} \cdot \cdots \cdot c_{i_n}.$$

Es ist also

$$\xi: \quad G_2 \to G_1$$

und

$$g_2\xi\alpha_1 = (h_2\eta\beta_1) \cdot c_{i_1} \cdot \cdots \cdot c_{i_n} = (h_2\beta_2) \cdot c_{i_1} \cdot \cdots \cdot c_{i_n} = g_2\alpha_2,$$

also

$$\xi\alpha_1 = \alpha_2,$$

d.h.

$$G_1(\alpha_1) \leqq G_2(\alpha_2).$$

Daraus folgen:

SATZ V. *Sind $H_1(\beta_1)$ und $H_2(\beta_2)$ verwandte Darstellungen eines direkten Faktors C von B, so sind die induzierten Darstellungen* (5) *von B auch verwandt.*

SATZ VI. *Induzieren die Darstellungen $H(\beta)$ und $H'(\beta')$ eines direkten Faktors C von B die Darstellungen $G(\alpha)$ und $G'(\alpha')$ von B, so induziert die Klasse der Darstellung $H(\beta)\,\tilde{\times}\,H'(\beta')$ (bzw. $H(\beta) * H'(\beta')$) die Klasse der Darstellung $G(\alpha)\,\tilde{\times}\,G'(\alpha')$ (bzw. $G(\alpha) * G'(\alpha')$) von B.*

Die Abbildung Φ, welche einer Darstellung $G(\alpha)$ von B die induzierte Darstellung $H(\beta)$ eines direkten Faktors C von B zuordnet, ist also eine verbandshomomorphe Abbildung von dem Pseudoverband $V(B)$ auf den Pseudoverband $V(C)$. Außerdem enthält $V(B)$ einen Teilpseudoverband $V^*(C)$, den man verbandsisomorph auf den Pseudoverband $V(C)$ eines direkten Faktors C von B abbilden kann (für jede direkte Komponente C von B) und derart, daß der Durchschnitt (bzw. Vereinigung) der Klassen in $V^*(C)$ mit dem entsprechenden Durchschnitt (bzw. Vereinigung) derselben Darstellungen in $V(B)$ übereinstimmt.

Es bleibt nun noch die Frage: Wie hängt der Pseudoverband $V(B)$ ab von den Pseudoverbänden $V(C_i)$ der direkten Faktoren C_i von B?

Es sei $G(\alpha)$ eine Darstellung von B und $H_i(\beta_i)$ die von $G(\alpha)$ induzierten Darstellungen der direkten Faktoren C_i von B; es gilt

$$H_i \cap H_j = K_\alpha \quad \text{für} \quad i \neq j,$$

wenn mit K_α der Kern der Abbildung α von G auf B bezeichnet wird. Wir bilden nun mit Hilfe der so gebildeten Darstellungen $H_i(\beta_i)$ das (verallgemeinerte) freie Produkt

$$F(\alpha) = \{\underset{i}{*} H_i(\beta_i) \,|\, K_\alpha\}(\alpha)$$

der $H_i(\beta_i)$ mit der vereinigten Untergruppe K_α. In diesem kann man die Elemente in einer Normalform schreiben. In jeder Gruppe $H_i(\beta_i)$ wählt man ein Repräsentantensystem $\{h_i\}$ der Nebenklassen modulo K_α; jedes Element $h \in H_i(\beta_i)$ hat dann eine eindeutige Darstellung $h = h_i \cdot k$ $(h_i \in \{h_i\},\ k \in K_\alpha)$. Dabei nehmen wir an, daß das Einselement $e_i \in \{h_i\}$, also K_α wird durch das Einselement repräsentiert. Es sind dann die Elemente des freien Produktes $F(\alpha)$ zu schreiben als

$$f = h_{i_1} \cdot h_{i_2} \cdot \cdots \cdot h_{i_n} \cdot k,$$

so daß

1. jede Komponente h_{i_j} $(1 \leqq j \leqq n)$ ein Element $(\neq e)$ eines Repräsentantensystems $\{h_{i_j}\}$ ist;
2. aufeinanderfolgende Komponenten h_{i_j} aus verschiedenen Repräsentantensystemen sind;
3. die letzte Komponente in der gemeinsamen Untergruppe K_α liegt;
4. $f\alpha = (h_{i_1} \cdot h_{i_2} \cdot \cdots \cdot h_{i_n} \cdot k)\alpha = (h_{i_1}\beta_{i_1}) \cdots (h_{i_n}\beta_{i_n})$.

Es läßt sich leicht beweisen, daß $F(\alpha)$ eine Darstellung von B ist, daß die Darstellung $F(\alpha)$ von B die Darstellungen $H_i(\beta_i)$ von C_i induziert, und daß außerdem $F(\gamma) \approx G(\alpha)$.

Wenn nun umgekehrt *irgendein* System von Darstellungen $H_i(\beta_i)$ von den direkten Faktoren C_i von B gegeben sei, so kann man die letzte Methode nicht ohne weiteres anwenden.

Wir bezeichnen mit K_{β_i} den Kern der Abbildung β_i von $H_i(\beta_i)$ auf C_i. In §1 (2.) haben wir schon bemerkt, daß mit $H_i(\beta_i)$ auch ein direktes Produkt

$$H_i'(\beta_i') = \{H_i(\beta_i) \times (\underset{j \neq i}{\times} K_{\beta_j})\}(\beta_i')$$

mit

$$(h_i k_{\beta_1} k_{\beta_2} \ldots k_{\beta_n})\beta_i' = h_i\beta_i \in C_i$$

eine Darstellung von C_i ist und außerdem, daß $H_i(\beta_i)$ und $H_i'(\beta_i')$ verwandte Darstellungen sind. Wenn wir nun für jede Darstellung $H_i(\beta_i)$ die entsprechende Darstellung $H_i'(\beta_i')$ bilden, so tritt für die $H_i'(\beta_i')$ genau der Fall auf, daß die $H_i'(\beta')$ Darstellungen von C_i sind und daß

$$H_i' \cap H_j' = \underset{i}{\times} K_{\beta_i} = K.$$

Wir können also auf analoge Weise das (verallgemeinerte) freie Produkt der $H_i'(\beta_i')$ mit der vereinigten Untergruppe $K = \underset{i}{\times} K_{\beta_i}$ bilden und es entsteht eine Darstellung $G(\alpha)$ von B, welche genau die $H_i'(\beta_i')$ als Darstellungen von den direkten Faktoren C_i von B induziert.

Da man jede Darstellung $G(\alpha)$ von B auf diese Weise bekommt, hat man folgendes Resultat:

SATZ VII. *Jede Darstellung $G(\alpha)$ von B bestimmt auf eindeutige Weise ein System von Darstellungen $H_i(\beta_i)$ der direkten Faktoren C_i von B. Umgekehrt kann man mit Hilfe irgendeines Systems von Darstellungen $H_i(\beta_i)$ der direkten*

Faktoren C_i von B durch Übergang auf gewisse verwandte Darstellungen $H_i'(\beta_i')$ von C_i erreichen, daß die entstehenden Darstellungen $H_i'(\beta_i')$ einen gemeinsamen Kern K besitzen. Das freie Produkt der $H_i'(\beta_i')$ mit der vereinigten Untergruppe K liefert dann eine Darstellung $G(\alpha)$ von B, die umgekehrt genau die Klassen der ursprünglichen Darstellungen $H_i(\beta_i)$ bestimmt.

Zum Schluß beweisen wir den

SATZ VIII. *Der Pseudoverband $V(B)$ ist das kardinalgeordnete Produkt der Pseudoverbände $V(C_i)$ der Faktoren C_i von B.*

BEWEIS. Es gibt eine eineindeutige Korrespondenz zwischen den Systemen von Klassen

$$[\ldots, \{H_i(\beta_i)\}, \ldots],$$

wo $H_i(\beta_i)$ eine Darstellung von C_i bezeichnet, und den Darstellungen

$$G(\alpha) = \{\underset{i}{\times} H_i\}(\alpha)$$

von B mit

$$g\alpha = (h_{i_1} \cdot \ldots \cdot h_{i_n})\alpha = (h_{i_1}\beta_{i_1}) \cdot \ldots \cdot (h_{i_n}\beta_{i_n}) \in B.$$

Wenn für die Komponentendarstellungen gilt $H_i(\beta_i) \leqq H_i'(\beta_i')$, so gilt auch für $G(\alpha) = (\underset{i}{\times} H_i)(\alpha)$ und $G'(\alpha') = (\underset{i}{\times} H_i')(\alpha')$:

$$G(\alpha) \leqq G'(\alpha')$$

und umgekehrt! Da man jede Darstellung $G(\alpha)$ von B auf diese Weise bekommt, ist $V(B)$ das kardinalgeordnete Produkt der Pseudoverbände $V(C_i)$ der direkten Faktoren C_i von B.

Technische Hogeschool, Delft

(Eingegangen am 11. April 1960)

ANDRÉ, J.
Math. Zeitschr. 76, 155–163 (1961)

Über Parallelstrukturen*)

Teil II: Translationsstrukturen

Von

JOHANNES ANDRÉ

Einleitung

Nachdem in Teil I dieser Arbeit[1]) die allgemeine Theorie der Parallelstrukturen entwickelt wurde, wenden wir uns jetzt speziellen Strukturen zu. Fordern wir von einer affinen Parallelstruktur, daß deren Translationsgruppe transitiv auf ihren Punkten ist, so kommt man zum Begriff der Translationsstruktur. Die Theorie dieser Strukturen wird ganz analog entwickelt wie in ANDRÉ (1954, a) [s. auch ANDRÉ (1954, b)] die der Translationsebenen (welche man kennzeichnen kann als Translationsstrukturen, die gleichzeitig affine Ebenen sind). Diejenigen Translationen, die eine feste Gerade in sich überführen, bilden eine Untergruppe der vollen Translationsgruppe T. Man betrachtet nun die zu jeder Geraden eines festen Büschels auf diese Weise gewonnene Untergruppe. Die Menge der so erhaltenen Untergruppen ist eine Partition von T, d.h. alle diese Untergruppen überdecken T, und je zwei verschiedene Untergruppen haben nur das neutrale Element gemeinsam. Umgekehrt erhält man aus einer Partition einer Gruppe eine Translationsstruktur, indem man als Punkte die Gruppenelemente und als Geraden die Rechtsrestklassen der Partitionenuntergruppen nimmt. Die Beweise hierfür stehen in §1.

Wichtig für die beiden weiteren Teile dieser Arbeit (zu deren Vorbereitung Teil II im wesentlichen dient) sind die Translationsstrukturen mit abelscher Translationsgruppe. Wie in §2 gezeigt wird, ist damit gleichbedeutend, daß jede Translation zentral ist, d.h. mit einer Geraden auch die zu ihr parallelen festläßt. Strukturen dieser Art nennen wir zentrale Translationsstrukturen (z.T.). Nach einer kleinen gruppentheoretischen Anwendung in §3 bringt §4 Untersuchungen über die Streckungsgruppe einer Translationsstruktur. Sie hängt eng zusammen mit der Menge aller derjenigen Endomorphismen der Translationsgruppe, welche jede Untergruppe der zur Struktur gehörenden Partition in sich abbildet; diese Menge heißt der Kern der Translationsstruktur. Die Streckungen bestehen aus genau denjenigen Kernelementen, die ein Inverses besitzen (Satz 4.2). Der Kern einer z.T. kann in natürlicher Weise als nullteilerfreier Ring mit Einselement aufgefaßt werden (Satz 4.3). Schließlich wird (in Satz 4.6) eine hinreichende Bedingung für einen Ring angegeben, damit er als Kern einer geeigneten z.T. aufgefaßt werden kann.

*) Teil einer von der Naturwissenschaftlich-Philosophischen Fakultät der Technischen Hochschule Braunschweig angenommenen Habilitationsschrift.

[1]) Teil I erschienen in Math. Z. **76**, 85–102 (1961), im folgenden hier mit I zitiert.

§ 1. Zusammenhang mit den Gruppenpartitionen

1. DEFINITION. Eine *Translationsstruktur* $t=(T, \mathfrak{L})$ (bezüglich einer Translationsgruppe T[2]) ist eine affine Parallelstruktur (a.P.) mit den beiden folgenden zusätzlichen Eigenschaften:

(t 1) *Es gibt eine auf den Punkten von t transitive Gruppe* T *von Translationen.*

(t 2) *Jede Gerade von t enthält mindestens zwei Punkte.*

BEMERKUNG. Das zweite Axiom ist eine Verschärfung des Reichhaltigkeitsaxiomes (a 4) für a.P.

Jede Gerade G einer Translationsstruktur ist Fixgerade mindestens einer nichtidentischen Translation. Nach (t 2) enthält nämlich G zwei Punkte p, q; es kann also $G=\langle p, q\rangle$ gesetzt werden. Aus (t 1) folgt die Existenz einer Translation $\tau(\in \mathsf{T})$ mit $p^\tau=q$. Wegen $G^\tau \| G$ folgt nun $G^\tau=G$. Allgemeinere Parallelstrukturen mit dieser Eigenschaft werden in Teil IV näher untersucht.

Es ist nicht bekannt, ob es eine Translationsstruktur t bezüglich mehrerer Translationsgruppen gibt, m.a.W. ob die Gruppe der axialen Kollineationen von $t=(T, \mathfrak{L})$ mehr als *eine* auf T transitive Translationsgruppe enthält. Dies ist jedenfalls nicht möglich, wenn alle Translationen von t eine Gruppe bilden, was sofort aus I, Satz 4.3 folgt. Offensichtlich gilt

SATZ 1.1. *Eine a.P. mit Axiom* (t 2), *in der die Translationen eine Gruppe* T *bilden, ist genau dann eine Translationsstruktur, wenn* T *transitiv auf ihren Punkten ist.*

Hieraus und aus I, Satz 5.3 (b) schließt man:

SATZ 1.2. *Eine endliche a.P. mit Axiom* (t 2) *ist genau dann eine Translationsstruktur, wenn die Gruppe* Π *ihrer affin-axialen Kollineationen transitiv auf ihren Punkten ist.*

2. Man gewinnt Translationsstrukturen durch sog. Gruppenpartitionen, ganz entsprechend wie man Translationsebenen durch Kongruenzen erhalten kann [ANDRÉ (1954, a), §§ 2, 3; PICKERT (1955), S. 199–201]. Sei T eine (abstrakte) Gruppe, deren Verknüpfung wir *additiv* schreiben und deren neutrales Element wir mit o bezeichnen wollen. Eine Menge $\mathfrak{P}$ aus *eigentlichen* (d.h. von o und T verschiedenen) Untergruppen von T nennen wir eine *Partition* [s. YOUNG (1927)] auf T, wenn die Untergruppen aus $\mathfrak{P}$ ganz T überdecken und je zwei verschiedene von ihnen nur das neutrale Element gemeinsam haben. Wir sagen, T sei die *Trägergruppe* (oder kurz der *Träger*) der Partition $\mathfrak{P}$. Jedes von o verschiedene $t \in T$ liegt in genau einer Untergruppe aus $\mathfrak{P}$, die wir mit U_t bezeichnen.

[2]) Wir benutzen hier eine gegenüber I etwas geänderte Bezeichnung: Hier ist T stets eine Gruppe, während in I mit T die Menge *aller* Translationen (von der nicht allgemein bekannt ist, ob sie eine Gruppe bildet) bezeichnet war. — Ferner werden hier mit Rücksicht auf die Teile III und IV die Punkt- bzw. Geradenmenge mit T bzw. $\mathfrak{L}$ anstatt mit P bzw. $\mathfrak{G}$ bezeichnet.

Wir können auf folgende Weise aus einer Partition $\mathfrak{P}$ einer Gruppe T eine Translationsstruktur $t=[T, \mathfrak{P}]$ gewinnen[3]). Als *Punkte* von t nehmen wir die Elemente von T, als *Geraden* die Rechtsrestklassen $U+t$ von Untergruppen $U \in \mathfrak{P}$ und als Translationen aus T die Abbildungen

$$\bar{t}: \quad x \to x+t \quad (x, t \in T).$$

Man zeigt nun ganz leicht, ganz entsprechend wie in ANDRÉ (1954, a), § 2, S. 166 [s. auch PICKERT (1955), S. 200], daß die so erhaltene Struktur eine a.P. ist, wenn man zwei Geraden $U+a$ und $V+b$ ($U, V \in \mathfrak{P}$; $a, b \in T$) genau dann *parallel* zueinander nennt, wenn $U=V$ ist: Durch die Punkte a, b geht die Gerade $U_{a-b}+b$ und keine weitere, die zu $U+a$ parallele Gerade durch b ist $U+b$; die Axiome (a 3) und (t 2) folgen daraus, daß $\mathfrak{P}$ aus eigentlichen Untergruppen besteht. Die Menge T aller Abbildungen $\bar{t}$ ist eine auf den Punkten von t transitive Gruppe von Translationen. Mithin ist t sogar eine Translationsstruktur. Übrigens ist die Abbildung $t \to \bar{t}$ ein Isomorphismus von T auf T.

3. Wir wollen jetzt umgekehrt zeigen, daß jede Translationsstruktur vermöge der oben beschriebenen Weise sich aus einer geeignet gewählten Gruppenpartition ergibt; m.a.W., daß jede Translationsstruktur isomorph zu einer Struktur $[T, \mathfrak{P}]$ ist. Der Beweis hierfür verläuft ganz analog wie der entsprechende in ANDRÉ (1954, a), §3.

Sei also t eine Translationsstruktur und T eine [nach (t 1) existierende] auf den Punkten von t transitive Translationsgruppe. Ferner sei o ein beliebiger, aber dann fester Punkt von t (der „Nullpunkt" oder „Ursprungspunkt"). Ist t ein weiterer Punkt, so gibt es genau eine Translation in T, die o in t überführt; sie werde $\bar{t}$ genannt. Die *Summe* t_1+t_2 zweier Punkte t_1 und t_2 (bezüglich T) werde definiert durch:

$$t_1 + t_2 = t_1^{\bar{t}_2} = o^{\bar{t}_1 \bar{t}_2}. \tag{1.1}$$

Auf diese Weise wird die Menge T der Punkte von t zu einer Gruppe, die durch die Zuordnung $t \to \bar{t}$ als zu T isomorph erkannt wird.

Sei nun U eine Gerade von t durch o. Gilt $U^{\bar{t}}=U$, so $o^{\bar{t}}=t \in U$. Sei umgekehrt $t \in U$. Dann gilt $o \in U \cap U^{\bar{t}^{-1}}$. Andererseits ist aber $U \parallel U^{\bar{t}^{-1}}$, weil $\bar{t}$ und damit auch $\bar{t}^{-1}$ eine Translation aus T ist. Aus diesen beiden Bemerkungen folgt $U^{\bar{t}^{-1}}=U$, also auch $U^{\bar{t}}=U$. Jetzt folgt sofort [wie in ANDRÉ (1954, a), S. 166], daß die Punkte aus U eine Untergruppe von T bilden (die wir ebenfalls U nennen wollen). Aus (a 3) folgt $U < T$, und aus (t 2) ergibt sich, daß U nicht nur aus o allein besteht; U ist also eigentliche Untergruppe von T. Alle so erhaltenen Untergruppen bilden eine (von t und T abhängige) Partition $\mathfrak{P}=[t, \mathsf{T}]$, weil jeder von o verschiedene Punkt von t auf genau einer Geraden durch o liegt. Durch t und T ist $[t, \mathsf{T}]$ eindeutig bestimmt, denn bei der soeben angegebenen Konstruktion war nur die Wahl des Nullpunktes o willkürlich;

[3]) Dieser Sachverhalt wurde bereits im Vortragsauszug von ANDRÉ (1954, b) angedeutet: Die Struktur $[T, \mathfrak{P}]$ wurde dort $j(\mathfrak{P})$ genannt.

es ist jedoch wegen der Transitivität von T gleichgültig, welchen Punkt von t man als Nullpunkt wählt. Die Translationen von T haben wegen (1.1) die Form $x \to x+t$; die Geraden von t ergeben sich mithin als Rechtsrestklassen von Untergruppen aus $\mathfrak{P}$. Damit ist die gewünschte Umkehrung vollständig bewiesen.

4. Vermöge der in den Ziffern 2 und 3 angegebenen Verfahren kann jeder Gruppenpartition eine Translationsstruktur und umgekehrt jeder Translationsstruktur eine Gruppenpartition zugeordnet werden. Durch Hintereinanderausführung dieser beiden Schritte gelangt man wieder zur Ausgangsstruktur, d.h. es gilt:

$$t = [T, [t, \mathsf{T}]], \qquad \mathfrak{P} = [[T, \mathfrak{P}], \mathsf{T}]. \tag{1.2}$$

[Hierbei muß für die Gültigkeit der ersten Gleichung von (1.2) angenommen werden, daß die Geraden von t als Punktmengen aufgefaßt werden, was wegen (t 2) möglich ist.] Die einfachen Beweise hierfür verlaufen ebenso wie die entsprechenden in ANDRÉ (1954, a), S. 166/67 und dürfen deshalb hier übergangen werden.

Zusammenfassend erhalten wir also den folgenden [dem Satz 9 von ANDRÉ (1954) entsprechenden]

SATZ 1.3 (HAUPTSATZ ÜBER TRANSLATIONSSTRUKTUREN). *Es gibt eine eineindeutige Zuordnung zwischen den Translationsstrukturen einerseits und den Gruppenpartitionen andererseits:*

Nimmt man die Elemente einer (abstrakten) Gruppe T als Punkte und die Rechtsrestklassen der Untergruppen einer Partition von T als Geraden, so erhält man eine Translationsstruktur t.

Umgekehrt erhält man aus jeder Translationsstruktur t eine Gruppenpartition, deren Trägergruppe (bis auf Isomorphie) eine auf den Punkten von t transitive Translationsgruppe T ist, und deren Untergruppen aus genau denjenigen Translationen von T bestehen, die je eine Gerade von t (als Ganzes) festlassen.

BEMERKUNG. Die von einer Partition $\mathfrak{P}$ erzeugte Translationsstruktur $[T, \mathfrak{P}]$ ist genau dann eine Translationsebene, wenn $T = U + V$ für irgend zwei verschiedene $U, V \in \mathfrak{P}$ gilt. In diesem Fall ist $\mathfrak{P}$ eine *Kongruenz* [vgl. ANDRÉ (1954, a), §2; PICKERT (1955), S. 200].

§ 2. Zentrale Translationsstrukturen

1. In dem vorigen Paragraphen wurde gezeigt, daß die Translationsstrukturen eng mit den Partitionen von Gruppen zusammenhängen. Zur genaueren Untersuchung der Translationsstrukturen können wir also gruppentheoretische Methoden heranziehen. Andererseits liefern geometrische Untersuchungen auch gruppentheoretische Ergebnisse (vgl. insbesondere §3). Wir wollen zunächst die zentralen Translationen [also die Translationen mit (uneigentlichem) Zentrum] einer Translationsstruktur kennzeichnen.

2. Die Bezeichnungen seien dieselben wie im vorigen Paragraphen. Diejenigen Translationen aus T, welche ein $U \in \mathfrak{P}$ in sich überführen, bilden eine Untergruppe $\mathsf{T}(U)$ von T; sie besteht aus allen Abbildungen $t \to t+u$ mit $u \in U$ (vgl. §1, Ziffer 3). Die Menge aller Translationen aus $\mathsf{T}(U)$ bildet also ein festes $t \in T$ auf die Linksrestklasse $t+U$ ab:

$$t^{\mathsf{T}(U)} = t+U = (t+U-t)+t. \tag{2.1}$$

Diese Restklasse ist genau dann eine Gerade, wenn mit U auch die zu U konjugierte Gruppe $t+U-t$ zu $\mathfrak{P}$ gehört, d.h.:

SATZ 2.1. *Genau dann ist jede Punktmenge $t^{\mathsf{T}(U)}$ $(t \in T,\ U \in \mathfrak{P})$, die durch Anwendung aller U festlassenden Translationen auf einen festen Punkt t entsteht, eine Gerade von $[T, \mathfrak{P}]$, wenn mit einer Untergruppe auch alle konjugierten zur Partition $\mathfrak{P}$ gehören.*

Auch eine Gerade der Form $t+U=(t+U-t)+t$ ist nicht notwendig parallel zu U. Parallelität gilt vielmehr dann und nur dann, wenn $t+U-t=U$ für alle $t \in T$ ist; das ergibt:

SATZ 2.2. *Die Translationen aus $\mathsf{T}(U)$ sind genau dann alle zentral, wenn U ein Normalteiler in T ist.*

Wir nennen eine Translationsstruktur *zentral* (zentrale Translationsstruktur, abgekürzt z.T.), wenn alle ihre Translationen zentral sind[4]). Da jede Translation in einem $\mathsf{T}(U)$ liegt, sind also genau dann alle Translationen zentral, wenn jedes $U \in \mathfrak{P}$ Normalteiler in T ist. Über Partitionen mit dieser Eigenschaft gibt Auskunft der

SATZ 2.3 [KONTOROWITSCH (1943)]. *Genau dann ist eine Gruppe T mit einer Partition $\mathfrak{P}$ abelsch, wenn alle Untergruppen von $\mathfrak{P}$ Normalteiler in T sind.*

BEWEIS. (a) T ist abelsch: Dann sind alle Untergruppen Normalteiler, insbesondere also die von $\mathfrak{P}$.

(b) Alle Untergruppen aus $\mathfrak{P}$ sind Normalteiler: Sei $t \in T$, $t \neq o$. Mit U_t bezeichnen wir wieder die (eindeutig bestimmte) Untergruppe von $\mathfrak{P}$, die t enthält. Es genügt, $t+t'=t'+t$ für $t, t', t+t' \neq o$ zu zeigen. Zunächst sei $U_t \neq U_{t'}$. Wegen $U_t \cap U_{t'} = \{o\}$ und der Normalteilereigenschaft von U_t und $U_{t'}$ in T, also auch in $U_t + U_{t'}$, können wir $U_t + U_{t'} = U_t \oplus U_{t'}$ als direkte Summe schreiben. Bei direkter Summation sind aber Elemente aus verschiedenen Summanden miteinander vertauschbar, insbesondere also auch t mit t'. Sei nun $U_t = U_{t'}$. Dann gibt es ein $t'' \in T$, das nicht in U_t liegt. Hieraus folgt $U_{t''} \neq U_{t'} = U_{t+t'}$ und $U_t \neq U_{t'+t''}$, und wir erhalten

$$t''+t+t' = t+t'+t'' = t'+t''+t = t''+t'+t,$$

also auch in diesem Falle $t+t'=t'+t$. Damit ist T als abelsch erkannt.

[4]) Nach I, Satz 4.5 (b) folgt daraus, daß die Menge aller Translationen eine Gruppe bilden (die mit T übereinstimmt).

Aus den Sätzen 2.2 und 2.3 folgt sofort der

SATZ 2.4. *Eine Translationsstruktur ist genau dann zentral, wenn ihre Translationsgruppe abelsch ist.*

BEMERKUNG. Hieraus folgt, daß nicht jede Translation (also auch nicht jede axiale Kollineation) zentral ist; denn es gibt Translationsstrukturen mit nichtabelschen Translationsgruppen, oder (was nach Satz 1.3 dasselbe besagt) nichtabelsche Gruppen mit Partitionen. Einfachste Beispiele solcher Gruppen sind die Diedergruppen D_{2n} [vgl. etwa SUZUKI (1950)]. Andere nichtabelsche Gruppen mit Partitionen sind die Frobenius-Gruppen Π mit $\Pi_p > \{1\}$ für $p \in P_1$ und $|P_1| > 1$ (zu den Bezeichnungen vgl. I, §5, insbesondere Satz 5.3). Denn die Untergruppen T und Π_p bilden in diesem Falle eine Partition von Π; die Π_p sind paarweise konjugiert, also keine Normalteiler in Π; nach Satz 2.3 ist also Π nicht abelsch.

§ 3. Eine weitere gruppentheoretische Anwendung

Wir wollen jetzt die Ergebnisse des §1 benutzen, um eine weitere Eigenschaft von Gruppen mit Partitionen herzuleiten, die ebenfalls bereits KONTOROWITSCH (1943) bewiesen hat. Zunächst zeigen wir:

SATZ 3.1. *Sei $|\mathfrak{P}|$ die (endliche oder unendliche) Mächtigkeit einer Partition $\mathfrak{P}$ von T und $|U|$ die von $U \in \mathfrak{P}$. Dann gilt*

$$|\mathfrak{P}| \geqq \operatorname*{Max}_{U \in \mathfrak{P}} |U| + 1 \tag{3.1}$$

unter der Voraussetzung, daß $\operatorname{Max}|U|$ existiert. Das Gleichheitszeichen gilt genau dann, wenn die zugehörige Translationsstruktur $[T, \mathfrak{P}]$ eine Gerade enthält, die alle zu ihr nichtparallelen Geraden schneidet.

Denn $|\mathfrak{P}|$ ist die Anzahl der Geraden von $t = [T, \mathfrak{P}]$ durch den Punkt o [und damit wegen (a 2) auch die Anzahl der Geraden durch einen beliebigen Punkt von t], und $|U|$ ist die Anzahl der Punkte auf jeder zu U parallelen Geraden L. Es gibt nun mindestens $|U|+1$ Geraden durch o, nämlich die $|U|$ Verbindungsgeraden von o mit den Punkten von $L \neq U$ und U selbst. Also gilt $|\mathfrak{P}| \geqq |U|+1$ und somit (3.1). Genau dann wird L von allen nicht zu L parallelen Geraden geschnitten, wenn $|\mathfrak{P}| = |U|+1$ ist. Damit ist Satz 3.1 vollständig bewiesen. Aus ihm folgt sofort der

ZUSATZ. *Gibt es in einer Gruppe T eine Partition mit nur endlich vielen Untergruppen, so ist T selbst endlich* [KONTOROWITSCH (1943)].

§ 4. Der Kern einer Gruppenpartition

1. Die Streckungen einer Translationsstruktur $t = [T, \mathfrak{P}]$ lassen sich ganz analog gewinnen wie die einer Translationsebene [vgl. ANDRÉ (1954, a), §4; PICKERT (1955), S. 202ff.]. Wie üblich ist ein *Endomorphismus* einer Gruppe T als homomorphe Abbildung von T in sich erklärt. Sei $\mathfrak{P}$ eine Partition von T. Ein *$\mathfrak{P}$-Endomorphismus* $\varkappa$ ist ein Endomorphismus von T, der jede Unter-

gruppe von $\mathfrak{P}$ in sich abbildet: *Aus* $U \in \mathfrak{P}$ *folgt* $U^{\varkappa} \subseteq U$. Für $\mathfrak{P}$-Endomorphismen $\varkappa, \lambda$ gilt:

(a) *Mit* $\varkappa, \lambda$ *ist auch* $\varkappa\lambda$ *ein* $\mathfrak{P}$*-Endomorphismus.*

(b) *Mit* $\varkappa$ *ist auch* $\varkappa^{-1}$ *(falls vorhanden) ein* $\mathfrak{P}$*-Endomorphismus.*

(c) *Die Endomorphismen* $\underline{0}$ *und* $\underline{1}$ (die durch $t^{\underline{0}} = o$ bzw. $t^{\underline{1}} = t$ für alle $t \in T$ definiert werden) *sind* $\mathfrak{P}$*-Endomorphismen.*

Die Eigenschaften (a) und (c) sind trivial; nur (b) bedarf noch eines Beweises: Sei $u \in U \in \mathfrak{P}$; wäre $u^{\varkappa^{-1}} \in V \in \mathfrak{P}$, $V \neq U$, so $u \in V^{\varkappa} \subseteq V$, $u = o$, also doch $u^{\varkappa^{-1}} (= o) \in U$. Aus der Existenz von $\varkappa^{-1}$ folgt also $U^{\varkappa^{-1}} \subseteq U$ für jedes $U \in \mathfrak{P}$, somit ist $\varkappa^{-1}$ ein $\mathfrak{P}$-Endomorphismus.

Die Menge aller $\mathfrak{P}$-Endomorphismen von T ist also hinsichtlich ihrer Multiplikation eine Halbgruppe mit Null- und Einselement; wir nennen diese Menge *Kern* von $\mathfrak{P}$ und bezeichnen sie mit $\mathsf{K}(\mathfrak{P})$ oder auch kurz mit K, wenn kein Mißverständnis zu befürchten ist.

Satz 4.1. *Jeder von* $\underline{0}$ *verschiedene* $\mathfrak{P}$*-Endomorphismus ist eine isomorphe Abbildung der Trägergruppe* T *in sich, d.h. ein Meromorphismus von* T [vgl. Zassenhaus (1937), S. 41].

Wir müssen dazu zeigen, daß $t^{\varkappa} \neq o$ aus $t \in T$, $t \neq o$ und $\varkappa \in \mathsf{K}$, $\varkappa \neq \underline{0}$ folgt. Der Beweis hierfür geht ganz entsprechend wie der Beweis (a) von Satz 10 aus André (1954, a) und darf deshalb übergangen werden. Das Produkt zweier Meromorphismen ist wieder ein Meromorphismus, also niemals $\underline{0}$, mithin gilt

Zusatz 1. *Die Halbgruppe* K *ist nullteilerfrei.*

Ist T *endlich*, so ist jede isomorphe Abbildung von T *in* sich sogar eine Abbildung *auf* sich, also jeder Meromorphismus ein Automorphismus; m.a.W.: aus $\varkappa \in \mathsf{K}$ und $\varkappa \neq \underline{0}$ folgt die Existenz von $\varkappa^{-1}$. In diesem Fall erhalten wir also den folgenden Satz 4.1 verschärfenden

Zusatz 2. *Die Menge* K^* *der von* $\underline{0}$ *verschiedenen* $\mathfrak{P}$*-Endomorphismen eines endlichen Trägers* T *ist eine Gruppe.*

Für unendliches T braucht dies nicht zu gelten, wie wir noch sehen werden (vgl. Ziffer 4). Ist aber $\mathfrak{P}$ eine Kongruenz, d.h. $[T, \mathfrak{P}]$ eine Translationsebene, so ist $\mathsf{K}^*(\mathfrak{P})$ stets eine Gruppe [vgl. André (1954), Satz 10].

2. Sei im allgemeinen Fall $\mathsf{K}^{\times}$ die Menge aller $\mathfrak{P}$-*Automorphismen* (d.h. der $\mathfrak{P}$-Endomorphismen, die zugleich Automorphismen sind). $\mathsf{K}^{\times}$ ist eine Gruppe. Jede Abbildung $\varkappa \in \mathsf{K}^{\times}$ ist eine Streckung in $t = [T, \mathfrak{P}]$ mit dem Fixpunkt o, und umgekehrt erhält man auf diese Weise alle o-Streckungen von t. Dies sieht man ebenso ein wie den entsprechenden Sachverhalt bei Translationsebenen [André (1954, a), S. 169], und man erhält den [zu André (1954, a), Satz 11 analogen[5])]

Satz 4.2. *Die Streckungen von* $t = [T, \mathfrak{P}]$ *mit dem Fixpunkt* o *(d.h. die* o*-Streckungen) sind genau die* $\mathfrak{P}$*-Automorphismen. Die Gruppen* $\mathsf{K}^{\times}$ *und* Π_o *stimmen überein.*

[5]) Die beiden dort angegebenen Zusätze gelten dagegen allgemein nur für endliches abelsches T; vgl. Ziffer 3.

BEMERKUNG. Aus Π_o gewinnt man leicht die übrigen Streckungsgruppen Π_t ($t \in T$); denn ist $\bar{t}$ die Translation, die o in t überführt, so gilt $\Pi_t = \bar{t}^{-1} \Pi_o \bar{t}$.

3. Wir setzen von nun an voraus, daß die Trägergruppe T *abelsch* sei. In diesem Falle ist die durch $t^{\varkappa+\lambda} = t^{\varkappa} + t^{\lambda}$ definierte *Summe* der $\mathfrak{P}$-Endomorphismen $\varkappa$ und λ wieder ein $\mathfrak{P}$-Endomorphismus. Auf diese Weise führen wir in dem Kern K neben der bereits definierten Multiplikation noch eine Addition ein; wir nennen die so gebildete Struktur ebenfalls Kern und bezeichnen sie wieder mit K, was wir ohne Befürchtung von Mißverständnissen tun können. Aus Zusatz 1 von Satz 4.1 folgt:

SATZ 4.3. *Der Kern* K *einer Partition* $\mathfrak{P}$ *mit abelschem Träger* T *ist (hinsichtlich der eingeführten Addition und Multiplikation) ein nullteilerfreier Ring mit Einselement. Man kann* T *auf natürliche Weise als* K*-Rechtsmodul auffassen (wobei das Einselement* $\underline{1}$ *von* K *auch Einheitsoperator von* T *ist).*

Sei $\mathfrak{P}$ eine Partition mit abelschem Träger T und dem Kern K, ferner $t = [T, \mathfrak{P}]$ die zugehörige Translationsstruktur (die nach Satz 2.4 zentral ist). Als *Charakteristik* von $\mathfrak{P}$ oder t bezeichnen wir die Charakteristik von K (d.h. die Ordnung von $\underline{1}$ hinsichtlich der Addition von K, wenn diese endlich ist, sonst 0). Die Charakteristik ist eine ganze nicht negative Zahl.

Für endliche abelsche Träger gilt wegen Zusatz 2 von Satz 4.1 schärfer:

SATZ 4.4. *Der Kern einer Partition* $\mathfrak{P}$ *mit endlichem abelschen Träger* T *ist ein Galois-Feld. Mithin ist* T *eine elementar abelsche Gruppe, d.h. alle Elemente* $t \neq o$ *von* T *haben die gleiche Ordnung* p*, die mit der Charakteristik von* $\mathfrak{P}$ *übereinstimmt und die in diesem Falle eine Primzahl ist.*

Jede endliche abelsche Trägergruppe ist also elementar abelsch. Ist andererseits T abelsch vom Typus $(p, \ldots, p)$ und nicht zyklisch, so bilden offenbar alle Untergruppen der Ordnung p eine (eigentliche) Partition. Also gilt:

SATZ 4.5 [YOUNG (1927)]. *Unter den endlichen abelschen Gruppen gestatten genau die nicht zyklischen elementar abelschen Gruppen eine (eigentliche) Partition.*

4. Für eine Kongruenz gilt über Satz 4.3 hinaus, daß ihr Kern stets ein Schiefkörper ist [ANDRÉ (1954, a), Satz 10]. Wie wir jetzt sehen werden, braucht dies für beliebige Partitionen mit abelscher Trägergruppe nicht mehr zu gelten. Vielmehr gilt:

SATZ 4.6. *Läßt sich ein Ring mit Einselement in einen Schiefkörper einbetten, so ist er (bis auf Isomorphie) Kern einer (geeignet gewählten) Partition einer abelschen Gruppe.*

BEWEIS. Es sei der Ring P mit Einselement in dem Schiefkörper $\overline{\mathsf{P}}$ eingebettet; ferner sei T ein Rechtsvektorraum über P mit $\dim T > 1$ und B eine Basis von T über P. Der Vektorraum $\overline{T} = B^{\overline{\mathsf{P}}}$ entsteht aus T durch Erweiterung von P zu $\overline{\mathsf{P}}$. Die Menge aller eindimensionalen $\overline{\mathsf{P}}$-Unterräume von $\overline{T}$ ist eine Partition $\overline{\mathfrak{P}}$ von $\overline{T}$, die Menge $\mathfrak{P}$ aller $\overline{U} \cap T > \{o\}$ mit $\overline{U} \in \overline{\mathfrak{P}}$ eine von T.

Wir wollen jetzt zeigen, daß der Kern K von $\mathfrak{P}$ mit P übereinstimmt. Jedenfalls gilt $\mathsf{P} \subseteqq \mathsf{K}$, denn die Elemente von $\mathfrak{P}$ sind P-Unterräume. Sei nun

$\varkappa \in \mathsf{K}$, o.B.d.A. sei $\varkappa \neq \underline{0}$. Das Basiselement b von T liege in der Partitionsuntergruppe $U_b = T \cap \overline{U}_b$ ($\overline{U}_b \in \overline{\mathfrak{B}}$). Es ist $U_b = b^{\mathsf{P}}$, wegen $b \in B$ und weil $\overline{U}_b$ über $\overline{\mathsf{P}}$ die Dimension 1 hat. Es gibt also ein $\varrho \in \mathsf{P}$, $\varrho \neq \underline{0}$ mit $b^{\varrho} = b^{\varkappa}$, also $b^{\varkappa\varrho^{-1}} = b$ ($\varrho^{-1} \in \overline{\mathsf{P}}$). Sei $c \notin U_b$, dann gilt $U_c \neq U_{b-c}$, also auch $\overline{U}_c \neq \overline{U}_{b-c}$. Aus $b - c^{\varkappa\varrho^{-1}} = (b-c)^{\varkappa\varrho^{-1}} \in \overline{U}_{b-c}$ folgt $c - c^{\varkappa\varrho^{-1}} \in \overline{U}_{b-c} \cap \overline{U}_c = \{o\}$, also $c = c^{\varkappa\varrho^{-1}}$. Ist $d \in U_b$, so folgt entsprechend wie eben (bei Ersetzung von b durch c), daß auch $d^{\varkappa\varrho^{-1}} = d$ ist. Daraus folgt $\varkappa\varrho^{-1} = \underline{1}$, $\varkappa = \varrho$, also $\mathsf{K} = \mathsf{P}$, was behauptet war.

Es ist unbekannt, ob Satz 4.6 die Ringe, welche Kerne einer Partition sein können, kennzeichnet. Wohl kann aber jetzt eine solche Charakterisierung für kommutative Ringe gegeben werden:

SATZ 4.7. *Genau dann ist ein kommutativer Ring isomorph zu einem Kern einer geeigneten Partition einer abelschen Gruppe, wenn er nullteilerfrei ist und ein Einselement hat.*

BEWEIS. Ein nullteilerfreier kommutativer Ring ist stets in seinen Quotientenkörper einbettbar. Hat er außerdem ein Einselement, so ist er nach Satz 4.6 als Kern einer Partition darstellbar. Der erste Teil von Satz 4.3 liefert die Umkehrung.

Literatur

ANDRÉ, J.: Über nicht-Desarguessche Ebenen mit transitiver Translationsgruppe. Math. Z. **60**, 156–186 (1954, a).

— Bemerkungen über Gruppenpartitionen. Proc. Intern. Math. Congr. Amsterdam **2**, 3 (1954, b).

KONTOROWITSCH, P.: Sur les groupes à base de partition. Mat. Sbornik **12**, 56–70 (1943).

PICKERT, G.: Projektive Ebenen. Berlin-Göttingen-Heidelberg: Springer 1955.

SUZUKI, M.: On the finite groups with a complete partition. J. Math. Soc. Jap. **2**, 165–185 (1950).

YOUNG, J. W.: On the partitions of a group and the resulting classification. Bull. Amer. Math. Soc. **33**, 453–461 (1927).

ZASSENHAUS, H.: Lehrbuch der Gruppentheorie, Bd. I. Leipzig: Teubner 1937.

Braunschweig, Mathematisches Institut der Technischen Hochschule

(Eingegangen am 13. Oktober 1960)

BABISTER, A. W.
Math. Zeitschr. 76, 164–173 (1961)

An inhomogeneous hypergeometric function

By

A. W. BABISTER

1. Introduction

In [1], the generalized modified Struve function $\Omega(a, c; x)$ was defined, this function being a particular integral of the inhomogeneous confluent hypergeometric equation

$$x\frac{d^2y}{dx^2}+(c-x)\frac{dy}{dx}-ay=\frac{2^{1-c}\Gamma(c)}{\Gamma(a)\Gamma(c-a)}e^{\frac{1}{2}x}.$$

In this paper, we consider the inhomogeneous hypergeometric function $B(a, b; c; z)$, defined in §3. $B(a, b; c; z)$ bears the same relation to the generalized modified Struve function Ω as $F(a, b; c; z)$ does to the confluent hypergeometric function Φ. As shown in §4, $B(a, b; c; z)$ is a particular integral of the equation

$$(1)\quad z(1-z)\frac{d^2u}{dz^2}+[c-(a+b+1)z]\frac{du}{dz}-abu=a\frac{2^{1-c}\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\left(1-\frac{1}{2}z\right)^{-a-1}.$$

2. Integral relations for the hypergeometric function

For ease of reference, we first summarize various integral relations for the solutions of the equation

$$(2)\qquad z(1-z)\frac{d^2u}{dz^2}+[c-(a+b+1)z]\frac{du}{dz}-abu=0,$$

where a, b and c are constants.

As shown in [3], the solutions of (2) can be expressed in terms of the simple loop integrals

$$(3)\qquad u_1=\frac{1}{2\pi i}\int_I^{(0+)}t^{b-1}(t-1)^{c-b-1}(zt-1)^{-a}dt,$$

$$(4)\qquad u_2=\frac{1}{2\pi i}\int_I^{(1+)}t^{b-1}(t-1)^{c-b-1}(zt-1)^{-a}dt,$$

and

$$(5)\qquad u_3=\frac{1}{2\pi i}\int_I^{(1/z+)}t^{b-1}(t-1)^{c-b-1}(zt-1)^{-a}dt,$$

where I is some convenient point in the t plane. If $\mathrm{re}\,(c-a)<2$, the integrands tend to zero as $|t|\to\infty$. Then, if

$$\arg z=\vartheta\qquad(-\pi<\vartheta<\pi),$$

we shall take I as the point at infinity having argument $\pi-\vartheta$, and

$$\arg t = \arg(t-1) = \arg(t-z^{-1}) = -\pi-\vartheta$$

at the start of the contour. The contours for u_1 and u_2 are shown in Fig. 1.

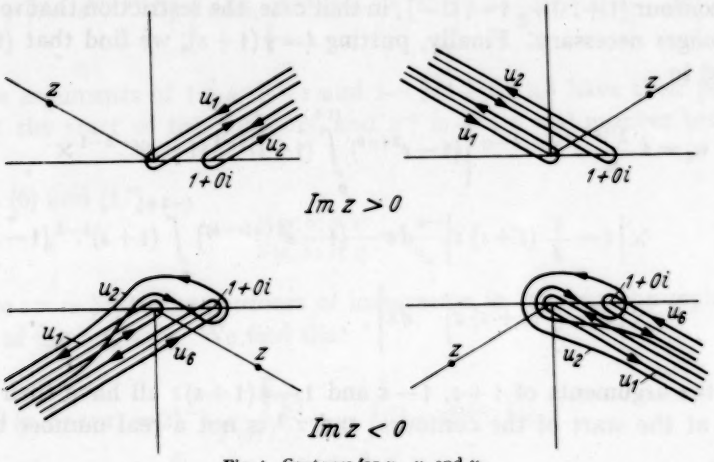

Fig. 1. Contours for u_1, u_2 and u_6

From [2], four solutions of (2) are

$$F(a,b;c;z) = i(2\pi)^{-1} e^{i\pi(b-a)} \Gamma(1-b)\,\Gamma(c)\,\Gamma(1+b-c)\,u_4, \tag{6}$$

$$z^{1-c} F(a-c+1, b-c+1; 2-c; z) = e^{-i\pi a}\Gamma(a)\,\Gamma(2-c)\,[\Gamma(a+1-c)]^{-1}\,u_5, \tag{7}$$

$$\begin{cases} F(a,b;a+b+1-c;1-z) \\ \quad = -e^{i\pi(c-a-b)}\,\Gamma(1-b)\,\Gamma(a+b+1-c)\,[\Gamma(a+1-c)]^{-1}\,u_1, \end{cases} \tag{8}$$

$$\begin{cases} z^{-a} F(a, a+1-c; a+1-b; z^{-1}) \\ = e^{i\pi(b+c-2a)}\,\Gamma(a+1-b)\,\Gamma(b+1-c)\,[\Gamma(a+1-c)]^{-1}\,u_2 \qquad (\operatorname{im} z > 0), \end{cases} \tag{9}$$

$$\begin{cases} z^{-a} F(a, a+1-c; a+1-b; z^{-1}) \\ = e^{i\pi(b-c)}\,\Gamma(a+1-b)\,\Gamma(b+1-c)\,[\Gamma(a+1-c)]^{-1}\,u_6 \qquad (\operatorname{im} z < 0), \end{cases} \tag{10}$$

where

$$\begin{cases} u_4 = \dfrac{1}{2\pi i} \displaystyle\int_I^{(1+,0+,1-,0-)} t^{b-1}(t-1)^{c-b-1}(zt-1)^{-a}\,dt \\ \quad = u_2\left[1-e^{2\pi i b}\right] - u_1\left[1-e^{2\pi i(c-b)}\right], \end{cases} \tag{11}$$

$$\begin{cases} u_5 = \dfrac{1}{2\pi i} \displaystyle\int_I^{(0+,1+)} t^{b-1}(t-1)^{c-b-1}(zt-1)^{-a}\,dt \\ \quad = u_1 + u_2\,e^{2\pi i b}, \end{cases} \tag{12}$$

and

$$u_6 = u_1[1-e^{2\pi i(c-b)}] + u_2 e^{2\pi i b}. \tag{13}$$

The contour for u_6 is shown in Fig. 1.

As shown in [2], the path of integration for u_4 may be contracted to a finite contour $(1+, 0+, 1-, 0-)$; in that case, the restriction that $\mathrm{re}\,(c-a)<2$ is no longer necessary. Finally, putting $t=\frac{1}{2}(1+s)$, we find that (11) can be reduced to

$$\left.\begin{aligned} u_4 = i\,\frac{2^{-c}}{\pi}\,e^{i\pi(a-b-c)}\Bigg[&(1-e^{2i\pi b})\int_0^{(1+)}(1+s)^{b-1}(1-s)^{c-b-1}\times \\ &\times\Big[1-\tfrac{1}{2}(1+s)\,z\Big]^{-a}\,ds - \{1-e^{2i\pi(c-b)}\}\int_0^{(-1+)}(1+s)^{b-1}(1-s)^{c-b-1}\times \\ &\times\Big[1-\tfrac{1}{2}(1+s)\,z\Big]^{-a}\,ds\Bigg], \end{aligned}\right\} \tag{14}$$

where the arguments of $1+s$, $1-s$ and $1-\frac{1}{2}(1+s)z$ all have their principal values at the start of the contours, and z^{-1} is not a real number between 0 and 1.

For $\mathrm{re}\,c>\mathrm{re}\,b>0$, the contours of integration in (14) may be replaced by portions of the real axis. We find that

$$\left.\begin{aligned} F(a,b;c;z) = 2^{1-c}\,\frac{\Gamma(c)}{\Gamma(b)\,\Gamma(c-b)}\int_0^1 (1-s^2)^{\frac{1}{2}c-1}\Big[\Big(\frac{1+s}{1-s}\Big)^{b-\frac{1}{2}c}\times \\ \times\Big[1-\tfrac{1}{2}(1+s)\,z\Big]^{-a} + \Big(\frac{1-s}{1+s}\Big)^{b-\frac{1}{2}c}\Big[1-\tfrac{1}{2}(1-s)\,z\Big]^{-a}\Big]\,ds. \end{aligned}\right\} \tag{15}$$

The corresponding formula for the confluent hypergeometric function can be deduced immediately, if we put $z=x/a$ and let $a\to\infty$. Then

$$\lim_{a\to\infty} F(a,b;c;x/a) = \Phi(b,c;x)$$

and (15) reduces to

$$\left.\begin{aligned} \Phi(b,c;x) = 2^{2-c}\,\frac{\Gamma(c)}{\Gamma(b)\,\Gamma(c-b)}\,e^{\frac{1}{2}x}\int_0^1 (1-s^2)^{\frac{1}{2}c-1}\times \\ \times\cosh\Big[\frac{1}{2}xs+\Big(b-\frac{1}{2}c\Big)\log\frac{1+s}{1-s}\Big]\,ds, \end{aligned}\right\} \tag{16}$$

in agreement with [1].

3. Inhomogeneous hypergeometric function $B(a, b; c; z)$

Consider the function $B(a,b;c;z)$ defined by

$$B(a,b;c;z) = i(2\pi)^{-1}\,e^{i\pi(b-a)}\,\Gamma(1-b)\,\Gamma(c)\,\Gamma(1+b-c)\,v, \tag{17}$$

where

$$(18)\quad \begin{cases} v = i\,\dfrac{2-c}{\pi}\,e^{i\pi(a-b-c)}\left[(1-e^{2i\pi b})\displaystyle\int_0^{(1+)}(1+s)^{b-1}(1-s)^{c-b-1}\left[1-\frac{1}{2}(1+s)z\right]^{-a}ds + \right. \\ \left. \quad +\{1-e^{2i\pi(c-b)}\}\displaystyle\int_0^{(-1+)}(1+s)^{b-1}(1-s)^{c-b-1}\left[1-\frac{1}{2}(1+s)z\right]^{-a}ds\right], \end{cases}$$

where the arguments of $1+s$, $1-s$ and $1-\frac{1}{2}(1+s)z$ all have their principal values at the start of the contours, and z^{-1} is not a real number between 0 and 1.

From (6) and (17),

$$\frac{B(a,b;c;z)}{F(a,b;c;z)} = \frac{v}{u_1}.$$

For re $c>$ re $b>0$, the contours of integration in (18) may be replaced by portions of the real axis. We find that

$$(19)\quad \begin{cases} B(a,b;c;z) = 2^{1-c}\,\dfrac{\Gamma(c)}{\Gamma(b)\,\Gamma(c-b)}\displaystyle\int_0^1(1-s^2)^{\frac{1}{2}c-1}\times \\ \quad \times\left[\left(\dfrac{1+s}{1-s}\right)^{b-\frac{1}{2}c}\left[1-\frac{1}{2}(1+s)z\right]^{-a}-\left(\frac{1-s}{1+s}\right)^{b-\frac{1}{2}c}\left[1-\frac{1}{2}(1-s)z\right]^{-a}\right]ds. \end{cases}$$

Put $z=x/a$ and let $a\to\infty$. The right hand side of (19) reduces to

$$2^{2-c}\,\frac{\Gamma(c)}{\Gamma(b)\,\Gamma(c-b)}\,e^{\frac{1}{2}x}\int_0^1(1-s^2)^{\frac{1}{2}c-1}\sinh\left[\frac{1}{2}xs+\left(b-\frac{1}{2}c\right)\log\frac{1+s}{1-s}\right]ds,$$

which, from [1], is equal to $\Omega(b,c;x)$, where Ω is the generalized modified Struve function. Thus

$$(20)\qquad \lim_{a\to\infty}B(a,b;c;x/a)=\Omega(b,c;x).$$

From [1], $\Omega(b,c;x)$ is a particular integral of the inhomogeneous confluent hypergeometric equation

$$(21)\qquad x\,\frac{d^2y}{dx^2}+(c-x)\,\frac{dy}{dx}-by=\frac{2^{1-c}\,\Gamma(c)}{\Gamma(b)\,\Gamma(c-b)}\,e^{\frac{1}{2}x}.$$

From (18), putting $t=\frac{1}{2}(1+s)$, we see that

$$(22)\quad \begin{cases} v=\dfrac{1}{2\pi i}\left[(1-e^{2i\pi b})\displaystyle\int_1^{(1+)}t^{b-1}(t-1)^{c-b-1}(zt-1)^{-a}dt + \right. \\ \left. \quad +\{1-e^{2i\pi(c-b)}\}\displaystyle\int_1^{(0+)}t^{b-1}(t-1)^{c-b-1}(zt-1)^{-a}dt\right], \end{cases}$$

where $\arg t = -2\pi$ and $\arg(t-1) = -\pi$ at the start of the contour. At that point, $\arg(t-z^{-1})$ is chosen so that it tends to $-\pi-\vartheta$ as z tends to zero (with argument ϑ).

Using (3), (4) and (13), we see, by deforming the contours, that, if $\mathrm{re}\,(c-a)<2$, (22) can be put in the form

$$(23)\quad \begin{cases} v = \dfrac{1}{2\pi i}\Big[(1-e^{2i\pi b})\Big\{\displaystyle\int_{\frac12}^{I} Z\,dt + 2\pi i\,u_2 - e^{2i\pi(c-b)}\int_{\frac12}^{I} Z\,dt\Big\} + \\ \qquad +\{1-e^{2i\pi(c-b)}\}\Big\{\displaystyle\int_{\frac12}^{I} Z\,dt + 2\pi i\,u_1 - e^{2i\pi b}\int_{\frac12}^{I} Z\,dt\Big\}\Big] \qquad (\mathrm{im}\,z>0), \end{cases}$$

$$(24)\quad \begin{cases} v = \dfrac{1}{2\pi i}\Big[(1-e^{2i\pi b})\Big\{\displaystyle\int_{\frac12}^{I} Z\,dt + 2\pi i\,e^{-2i\pi b}u_6 - e^{2i\pi(c-b)}\int_{\frac12}^{I} Z\,dt\Big\} + \\ +\{1-e^{2i\pi(c-b)}\}\Big\{\displaystyle\int_{\frac12}^{I} Z\,dt + 2\pi i\,e^{-2i\pi b}u_1 - e^{2i\pi b}\int_{\frac12}^{I} Z\,dt\Big\}\Big] \qquad (\mathrm{im}\,z<0), \end{cases}$$

where

$$Z = t^{b-1}(t-1)^{c-b-1}(zt-1)^{-a}$$

and the contours for u_1, u_2 and u_6 are given above.

The integrals $\int_{\frac12}^{I} Z\,dt$ can be taken along a line through $t=\frac12$ having argument $(\pi-\vartheta)$, where $\arg z=\vartheta$ $(-\pi<\vartheta<\pi)$. Putting $t=r+\frac12$, we find that

$$(25)\quad \int_{\frac12}^{I} Z\,dt = e^{2i\pi(a-c)}\int_{0}^{\infty\exp(i\pi-i\vartheta)} (r+\tfrac12)^{b-1}(r-\tfrac12)^{c-b-1}\{z(r+\tfrac12)-1\}^{-a}\,dr,$$

where the arguments of $r+\frac12$ and $r-\frac12$ are to be taken as 0 and π at the origin, and $\arg(r+\frac12-z^{-1})$ is chosen so that, at the origin, it tends to $\pi-\vartheta$ as z tends to zero (with argument ϑ).

From (17), (23) and (24), using (8)—(10), we find that

$$(26)\quad \begin{cases} B(a,b;c;z) = e^{i\pi(a+b-c)\varepsilon}\dfrac{\Gamma(c)\Gamma(a+1-c)}{\Gamma(b)\Gamma(a+1-b)}\,z^{-a}F(a,a+1-c;a+1-b;z^{-1}) - \\ \qquad - e^{i\pi b\varepsilon}\dfrac{\Gamma(c)\Gamma(a+1-c)}{\Gamma(c-b)\Gamma(a+b+1-c)}F(a,b;a+b+1-c;1-z) - \\ \qquad - 2e^{i\pi(a+b-c)}\dfrac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\displaystyle\int_{0}^{\infty\exp(i\pi-i\vartheta)}\Big(r+\frac12\Big)^{b-1}\Big(r-\frac12\Big)^{c-b-1}\times \\ \qquad \times\Big\{z\Big(r+\frac12\Big)-1\Big\}^{-a}\,dr \qquad [\mathrm{re}\,(c-a)<2], \end{cases}$$

where $\varepsilon=\mathrm{sgn}\,(\mathrm{im}\,z)=1,\ -1$ according as $\mathrm{im}\,z>0,\ <0$.

4. Differential equation satisfied by $B(a, b; c; z)$

As in [(2) 60], it can be shown that

$$(27)\quad \begin{cases} \left\{z(1-z)\frac{\partial^2}{\partial z^2}+[c-(a+b+1)z]\frac{\partial}{\partial z}-ab\right\}\times \\ \quad\times\left[(1+s)^{b-1}(1-s)^{c-b-1}\left\{1-\frac{1}{2}(1+s)z\right\}^{-a}\right] \\ =-\frac{1}{2}a\frac{\partial}{\partial s}\left[(1+s)^{b}(1-s)^{c-b}\left\{1-\frac{1}{2}(1+s)z\right\}^{-a-1}\right]. \end{cases}$$

From (17), (18) and (27), after some reduction, we find that $B(a, b; c; z)$ satisfies the inhomogeneous hypergeometric equation

$$(28)\quad z(1-z)\frac{d^2u}{dz^2}+[c-(a+b+1)z]\frac{du}{dz}-abu=a\frac{2^{1-c}\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\left(1-\frac{1}{2}z\right)^{-a-1}.$$

Equation (28) can be written in the alternative form

$$(29)\quad [\delta(\delta+c-1)-z(\delta+a)(\delta+b)]u=a\frac{2^{1-c}\Gamma(c)}{\Gamma(b)\Gamma(c-b)}z\left(1-\frac{1}{2}z\right)^{-a-1},$$

where

$$\delta\equiv z\frac{d}{dz}.$$

Differentiating (18) with respect to z, we obtain

$$(30)\quad \frac{dB(a, b; c; z)}{dz}=\frac{ab}{c}B(a+1, b+1; c+1; z),$$

$$(31)\quad \frac{d^nB(a, b; c; z)}{dz^n}=\frac{\Gamma(a+n)\Gamma(b+n)\Gamma(c)}{\Gamma(a)\Gamma(b)\Gamma(c+n)}B(a+n, b+n; c+n; z).$$

5. Relations for contiguous functions

From (6), (14), (17) and (18),

$$(32)\quad B(a, b; c; z)+F(a, b; c; z)=\frac{i}{\pi}2^{1-c}e^{i\pi(b-c)}\frac{\Gamma(c)\Gamma(1+b-c)}{\Gamma(b)}I(a, b; c; z),$$

where

$$(33)\quad I(a, b; c; z)=\int_0^{(1+)}(1+s)^{b-1}(1-s)^{c-b-1}[1-\tfrac{1}{2}(1+s)z]^{-a}ds.$$

From (33),

$$(34)\quad I(a+1; b; c; z)=I(a, b; c; z)+\tfrac{1}{2}zI(a+1, b+1; c+1; z),$$

and

$$(35)\quad 2I(a, b; c-1; z)=I(a, b; c; z)+I(a, b+1; c; z).$$

Integrating (33) by parts we obtain

$$\begin{aligned} bI(a, b; c; z)=&\{e^{2i\pi(c-b)}-1\}(1-\tfrac{1}{2}z)^{-a}+ \\ &+(c-b-1)\int_0^{(1+)}(1+s)^{b}(1-s)^{c-b-2}[1-\tfrac{1}{2}(1+s)z]^{-a}ds- \\ &-\tfrac{1}{2}az\int_0^{(1+)}(1+s)^{b}(1-s)^{c-b-1}[1-\tfrac{1}{2}(1+s)z]^{-a-1}ds, \end{aligned}$$

that is,

$$(36)\quad \begin{cases} b\,I(a,b;c;z) + (b-c+1)\,I(a,b+1;c;z) + \\ \quad + \frac{1}{2}a\,z\,I(a+1,b+1;c+1;z) = \{e^{2i\pi(c-b)}-1\}(1-\frac{1}{2}z)^{-a}. \end{cases}$$

From (34)—(36), using (32) and the recurrence formulae for F given in [2], we obtain

$$(37)\quad B(a+1,b;c;z) = B(a,b;c;z) + b\,c^{-1}\,z\,B(a+1,b+1;c+1;z),$$

$$(38)\quad (c-1)\,B(a,b;c-1;z) = (c-b-1)\,B(a,b;c;z) + b\,B(a,b+1;c;z),$$

$$(39)\quad \begin{cases} B(a,b+1;c;z) = B(a,b;c;z) + a\,c^{-1}\,z\,B(a+1,b+1;c+1;z) + \\ \quad + \dfrac{2^{2-c}\,\Gamma(c)}{\Gamma(b+1)\,\Gamma(c-b)}\left(1-\frac{1}{2}z\right)^{-a}. \end{cases}$$

We shall now show that the six contiguous functions $B(a\pm1,b;c;z)$, $B(a,b\pm1;c;z)$, $B(a,b;c\pm1;z)$ can be expressed in terms of $B(a,b;c;z)$ and its derivative with respect to z [cf. [3] 92]. As in [3], for conciseness, we denote these contiguous functions by B_{a+}, B_{a-}, etc.

From (30) and (37),

$$(40)\quad a\,B_{a+} = (\delta+a)\,B,$$

where $\delta \equiv z\frac{d}{dz}$. The operator δ is used in this sense in the following analysis. Now, from (29),

$$[\delta(\delta+c-1) - z(\delta+a-1)(\delta+b)]\,B_{a-} = (a-1)\,\frac{2^{1-c}\,\Gamma(c)}{\Gamma(b)\,\Gamma(c-b)}\,z\left(1-\frac{1}{2}z\right)^{-a},$$

that is,

$$(41)\quad \begin{cases} [\delta+c-a-z(\delta+b)](\delta+a-1)\,B_{a-} \\ \quad = (a-1)\,\dfrac{2^{1-c}\,\Gamma(c)}{\Gamma(b)\,\Gamma(c-b)}\,z\left(1-\frac{1}{2}z\right)^{-a} + (c-a)(a-1)\,B_{a-}. \end{cases}$$

From (40), with $a-1$ instead of a, and (41),

$$(42)\quad (c-a)\,B_{a-} = [(1-z)\,\delta + c - a - b\,z]\,B - \frac{2^{1-c}\,\Gamma(c)}{\Gamma(b)\,\Gamma(c-b)}\,z\left(1-\frac{1}{2}z\right)^{-a}.$$

From (30) and (39),

$$(43)\quad b\,B_{b+} = (\delta+b)\,B + \frac{2^{2-c}\,\Gamma(c)}{\Gamma(b)\,\Gamma(c-b)}\left(1-\frac{1}{2}z\right)^{-a}.$$

Now, from (29),

$$[\delta(\delta+c-1) - z(\delta+a)(\delta+b-1)]\,B_{b-} = a\,\frac{2^{1-c}\,\Gamma(c)}{\Gamma(b-1)\,\Gamma(c-b+1)}\,z\left(1-\frac{1}{2}z\right)^{-a-1},$$

that is,

$$(44)\quad \begin{cases} [\delta+c-b-z(\delta+a)](\delta+b-1)\,B_{b-} \\ \quad = a\,\dfrac{2^{1-c}\,\Gamma(c)}{\Gamma(b-1)\,\Gamma(c-b+1)}\,z\left(1-\frac{1}{2}z\right)^{-a-1} + (c-b)(b-1)\,B_{b-}. \end{cases}$$

From (43), with $b-1$ instead of b, and (44),

$$(45)\quad (c-b)\,B_{b-} = [(1-z)\,\delta + c - b - a\,z]\,B - \frac{2^{2-c}\,\Gamma(c)}{\Gamma(b)\,\Gamma(c-b)}\left(1-\frac{1}{2}z\right)^{-a}.$$

From (38) and (43),

$$(46)\qquad (c-1)\,B_{c-} = (\delta + c - 1)\,B + \frac{2^{2-c}\,\Gamma(c)}{\Gamma(b)\,\Gamma(c-b)}\left(1-\frac{1}{2}z\right)^{-a}.$$

Now, from (29),

$$[\delta(\delta+c) - z(\delta+a)(\delta+b)]\,B_{c+} = a\,\frac{2^{-c}\,\Gamma(c+1)}{\Gamma(b)\,\Gamma(c+1-b)}\,z\left(1-\frac{1}{2}z\right)^{-a-1},$$

that is,

$$(47)\quad \begin{cases} [\delta - z(\delta+a+b-c)]\,(\delta+c)\,B_{c+} \\ \qquad = a\,\dfrac{2^{-c}\,\Gamma(c+1)}{\Gamma(b)\,\Gamma(c+1-b)}\,z\left(1-\frac{1}{2}z\right)^{-a-1} + (c-a)\,(c-b)\,z\,B_{c+}. \end{cases}$$

From (46), with $c+1$ instead of c, and (47),

$$(48)\quad \begin{cases} (c-a)\,(c-b)\,z\,B_{c+} \\ \qquad = c\,[(1-z)\,\delta + z(c-a-b)]\,B - \dfrac{2^{1-c}\,\Gamma(c+1)}{\Gamma(b)\,\Gamma(c-b)}\,z\left(1-\frac{1}{2}z\right)^{-a}. \end{cases}$$

Equations (40), (42), (43), (45), (46) and (48) may be rearranged to give

$$(49)\qquad \delta B = a\,(B_{a+} - B),$$

$$(50)\qquad \delta B = b\,(B_{b+} - B) - \frac{2^{2-c}\,\Gamma(c)}{\Gamma(b)\,\Gamma(c-b)}\left(1-\frac{1}{2}z\right)^{-a},$$

$$(51)\qquad \delta B = (c-1)\,(B_{c-} - B) - \frac{2^{2-c}\,\Gamma(c)}{\Gamma(b)\,\Gamma(c-b)}\left(1-\frac{1}{2}z\right)^{-a},$$

$$(52)\quad (1-z)\,\delta B = (c-a)\,B_{a-} + (a-c+b\,z)\,B + \frac{2^{1-c}\,\Gamma(c)}{\Gamma(b)\,\Gamma(c-b)}\,z\left(1-\frac{1}{2}z\right)^{-a},$$

$$(53)\quad (1-z)\,\delta B = (c-b)\,B_{b-} + (b-c+a\,z)\,B + \frac{2^{2-c}\,\Gamma(c)}{\Gamma(b)\,\Gamma(c-b)}\left(1-\frac{1}{2}z\right)^{-a},$$

$$(54)\quad \begin{cases} c\,(1-z)\,\delta B = (c-a)\,(c-b)\,z\,B_{c+} + c\,(a+b-c)\,z\,B + \\ \qquad + \dfrac{2^{1-c}\,\Gamma(c+1)}{\Gamma(b)\,\Gamma(c-b)}\,z\left(1-\frac{1}{2}z\right)^{-a}. \end{cases}$$

The 15 relations between the six contiguous functions (corresponding to Gauss's relations for the hypergeometric function) follow from (49)—(54), on equating any two values of δB.

6. Associated differential equations

From (17) and (18),

$$(55)\qquad B(a,b;c;z) = -(1-z)^{-a}\,B[a,c-b;c;z/(z-1)].$$

Thus $(1-z)^{-a}\,B[a,b;c;z/(z-1)]$ is a particular integral of the equation

$$z(1-z)\frac{d^2u}{dz^2} + [c-(a+b+1)\,z]\frac{du}{dz} - a\,b\,u = -a\,\frac{2^{1-c}\,\Gamma(c)}{\Gamma(b)\,\Gamma(c-b)}\left(1-\frac{1}{2}z\right)^{-a-1}.$$

B function	$f(z)$
$B(a, b; c; z)$ $(1-z)^{-a} B[a, c-b; c; z/(z-1)]$	$\pm a \dfrac{2^{a-c+2}\Gamma(c)}{\Gamma(b)\Gamma(c-b)} (2-z)^{-a-1}$
$(1-z)^{c-a-b} B(c-a, c-b; c; z)$ $(1-z)^{-b} B[c-a, b; c; z/(z-1)]$	$\pm (c-a) \dfrac{2^{2-c}\Gamma(c)}{\Gamma(c-b)\Gamma(b)} \times$ $\times (2-z)^{a-c-1} (1-z)^{c-a-b}$
$z^{1-c} B(b+1-c, a+1-c; 2-c; z)$ $z^{1-c}(1-z)^{c-b-1} \times$ $\times B[b+1-c, 1-a; 2-c; z/(z-1)]$	$\pm (b+1-c) \dfrac{2^{b+1}\Gamma(2-c)}{\Gamma(a+1-c)\Gamma(1-a)} \times$ $\times (2-z)^{c-b-2} z^{1-c}$
$z^{1-c}(1-z)^{c-a-b} B(1-b, 1-a; 2-c; z)$ $z^{1-c}(1-z)^{c-a-1} \times$ $\times B[1-b, a+1-c; 2-c; z/(z-1)]$	$\pm (1-b) \dfrac{2^{c-b+1}\Gamma(2-c)}{\Gamma(1-a)\Gamma(a+1-c)} \times$ $\times (2-z)^{b-2} (1-z)^{c-a-b} z^{1-c}$
$B(a, b; a+b+1-c; 1-z)$ $z^{-a} B(a, a+1-c; a+b+1-c; 1-z^{-1})$	$\pm a \dfrac{2^{c-b+1}\Gamma(a+b+1-c)}{\Gamma(b)\Gamma(a+1-c)} (1+z)^{-a-1}$
$z^{1-c} B(b+1-c, a+1-c;$ $a+b+1-c; 1-z)$ $z^{-b} B(b+1-c, b; a+b+1-c; 1-z^{-1})$	$\pm (b+1-c) \dfrac{2^{2-a}\Gamma(a+b+1-c)}{\Gamma(a+1-c)\Gamma(b)} \times$ $\times (1+z)^{c-b-2} z^{1-c}$
$(1-z)^{c-a-b} B(c-a, c-b;$ $c+1-a-b; 1-z)$ $z^{a-c}(1-z)^{c-a-b} B(c-a, 1-a;$ $c+1-a-b; 1-z^{-1})$	$\pm (c-a) \dfrac{2^{b+1}\Gamma(c+1-a-b)}{\Gamma(c-b)\Gamma(1-a)} \times$ $\times (1+z)^{a-c-1} (1-z)^{c-a-b}$
$z^{1-c}(1-z)^{c-a-b} \times$ $\times B(1-b, 1-a; c+1-a-b; 1-z)$ $z^{b-c}(1-z)^{c-a-b} \times$ $\times B(1-b, c-b; c+1-a-b; 1-z^{-1})$	$\pm (1-b) \dfrac{2^{a-c+2}\Gamma(c+1-a-b)}{\Gamma(1-a)\Gamma(c-b)} \times$ $\times (1+z)^{b-2} (1-z)^{c-a-b} z^{1-c}$
$(-z)^{-a} B(a, a+1-c; a+1-b; z^{-1})$ $(1-z)^{-a} B[a, c-b; a+1-b; (1-z)^{-1}]$	$\pm a \dfrac{2^{b+1}\Gamma(a+1-b)}{\Gamma(a+1-c)\Gamma(c-b)} (1-2z)^{-a-1}$
$(-z)^{b-c}(1-z)^{c-a-b} \times$ $\times B(1-b, c-b; a+1-b; z^{-1})$ $(-z)^{1-c}(1-z)^{c-a-1} \times$ $\times B[1-b, a+1-c; a+1-b; (1-z)^{-1}]$	$\pm (1-b) \dfrac{2^{2-a}\Gamma(a+1-b)}{\Gamma(c-b)\Gamma(a+1-c)} \times$ $\times (1-2z)^{b-2} (1-z)^{c-a-b} (-z)^{1-c}$
$(-z)^{-b} B(b+1-c, b; b+1-a; z^{-1})$ $(-z)^{1-c}(1-z)^{c-b-1} \times$ $\times B[b+1-c, 1-a; b+1-a; (1-z)^{-1}]$	$\pm (b+1-c) \dfrac{2^{a-c+2}\Gamma(b+1-a)}{\Gamma(b)\Gamma(1-a)} \times$ $\times (1-2z)^{c-b-2} (-z)^{1-c}$
$(-z)^{a-c}(1-z)^{c-a-b} \times$ $\times B(c-a, 1-a; b+1-a; z^{-1})$ $(1-z)^{-b} B[c-a, b; b+1-a; (1-z)^{-1}]$	$\pm (c-a) \dfrac{2^{c-b+1}\Gamma(b+1-a)}{\Gamma(1-a)\Gamma(b)} \times$ $\times (1-2z)^{a-c-1} (1-z)^{c-a-b}$

There are 24 functions of the type

$$z^{\lambda}(1-z)^{\mu} B(a', b'; c'; z')$$

which correspond to Kummer's 24 solutions of the hypergeometric equation. These functions are particular integrals of equations of the form

$$z(1-z)\frac{d^2u}{dz^2} + [c-(a+b+1)z]\frac{du}{dz} - abu = f(z).$$

The 24 functions, grouped in pairs according to (55), are given in the table, together with the corresponding values of $f(z)$.

References

[1] Babister, A. W.: Generalized modified Struve functions. Quart. J. Math., Oxford (2) **10**, 214–223 (1959).

[2] Erdélyi, A.: Higher transcendental functions, vol. 1. Bateman Manuscript Project, 1953.

[3] Poole, E. G. C.: Linear differential equations. Oxford 1936.

The University, Glasgow

(Eingegangen am 25. Juli 1960)

Berz, E.
Math. Zeitschr. 76, 174–198 (1961)

Lösung der linearen Differentialgleichung durch Transformation in einen Schiefkörper

Von

Edgar Berz

Einleitung

Sind die Koeffizienten der linearen Differentialgleichung

$$\text{(I)} \qquad Y^{(n)}(t) + a_1 Y^{(n-1)}(t) + \cdots + a_n Y(t) = F(t)$$

konstant, so kann man die Gleichung bei geeignetem $F(t)$ mittels der Laplace-Transformation $\mathfrak{L}$ lösen. Das Prinzip ist dabei folgendes: Man wendet $\mathfrak{L}$ auf (I) an. Dies ergibt für das Bild $y(s) = \mathfrak{L}\{Y(t)\}$ eine einfache algebraische Gleichung

$$a(s)\, y(s) = b(s).$$

Zur Bildfunktion $y = a^{-1} b$ ist dann noch die Urfunktion $Y(t)$ zu bestimmen. Diese ist Lösung von (I).

Eine ganz entsprechende Methode wird in dieser Arbeit für den Fall entwickelt, daß die Koeffizienten der Differentialgleichung sowie F [1]) analytische Funktionen sind, etwa um $t = 0$.

An die Stelle der Laplace-Transformation tritt dabei eine Transformation $\mathfrak{T}$, die für die Klasse C_δ der in

$$|s| < \delta, \qquad |t| < \delta$$

analytischen Funktionen $K(s, t)$ erklärt ist. Der Bildbereich ist nicht wie oben eine Funktionenklasse, sondern ein Schiefkörper SK. Durch diese Transformation läßt sich die Differentialgleichung in eine algebraische Gleichung $a x = b$ in bezug auf den Schiefkörper SK überführen. Zu der Lösung $x = a^{-1} b$ im Bildbereich SK läßt sich dann eine Urfunktion bestimmen und diese ist Lösung der vorgelegten Differentialgleichung.

§ 1. Der Operatorenring $\mathsf{K}[s, I]$

Sei K der Körper der komplexen Zahlen und $\mathfrak{C}_\delta$ die Klasse aller komplexen Funktionen $f(s)$, die für $|s| < \delta$ regulär sind. Eine Vorschrift A, die jedem $f \in \mathfrak{C}_\delta$ in eindeutiger Weise ein Element Af zuordnet, heiße Operator in $\mathfrak{C}_\delta$. Für diese Operatoren $A, B, \ldots$ definieren wir Gleichheit, Summe und Produkt in üblicher Weise:

$$A = B \text{ genau dann, wenn } Af = Bf \quad \text{für alle } f \in \mathfrak{C}_\delta,$$
$$(AB) f = A(Bf), \qquad (A + B) f = Af + Bf.$$

[1]) Die Resultate sind bei Beschränkung auf das Reelle auch für stetiges F gültig.

A heiße additiv, wenn für jedes Paar $f, g \in \mathfrak{C}_\delta$ gilt:

$$A(f+g) = Af + Ag.$$

Die Gesamtheit dieser additiven Operatoren in $\mathfrak{C}_\delta$ bilden offenbar einen Ring. Aus diesem greifen wir folgende Elemente heraus:

1. Multiplikation mit einer Konstanten:

Sei c eine Konstante aus K und es werde $f(s)$ in $c \cdot f(s)$ abgebildet. Diesen Operator bezeichnen wir ebenfalls mit c:

$$cf = c \cdot f.$$

2. Multiplikation mit s:

Sei s die komplexe Veränderliche und es werde $f(s)$ in $s \cdot f(s)$ überführt. Dieser Operator wird mit s bezeichnet:

$$sf = s \cdot f.$$

3. Integrationsoperator:

Die Zuordnung

$$f(s) \to \int_0^s f(\sigma)\, d\sigma \quad \text{mit} \quad |s| < \delta$$

werde als „Integrationsoperator I" eingeführt.

Diese Operatoren 1. bis 3. erzeugen für sich einen Ring. Er besteht aus allen Operatoren, die sich aus 1. bis 3. durch endliche Verknüpfung (Addition und Multiplikation) ergeben. Diesen Ring bezeichnen wir durch $\mathsf{K}[s, I]$.

Wir wollen nun die Struktur von $\mathsf{K}[s, I]$ angeben. Hierfür ist grundlegend die Vertauschungsrelation:

$$Is = sI - I^2.$$

Man beweist sie leicht durch partielle Integration:

$$(Is)f = \int_0^s \sigma f(\sigma)\, d\sigma = s\int_0^s f(\sigma)\, d\sigma - \int_0^s \left(\int_0^\sigma f(\sigma')\, d\sigma'\right) d\sigma = (sI - I^2)f.$$

Wendet man diese Vertauschungsrelation konsequent auf jedes Produkt Is an, so erhält man nur „geordnete" Produkte, bei denen s vor I kommt. Es läßt sich demnach jedes Element aus $\mathsf{K}[s, I]$ als „geordnetes Polynom"

$$f(s, I) = \sum_{i,k} c_{ik}\, s^i I^k$$

mit Koeffizienten aus K darstellen.

Ohne große Mühe läßt sich beweisen, daß diese Darstellung durch ein geordnetes Polynom eindeutig ist.

Man kann daher die Elemente von $\mathsf{K}[s, I]$ umkehrbar eindeutig den gewöhnlichen Polynomen $f(u, v^{-1})$ in u, v^{-1} zuordnen vermöge:

$$f(s, I) = \sum c_{ik}\, s^i I^k \to \sum c_{ik}\, u^i v^{-k} = f(u, v^{-1}).$$

Diese Zuordnung ist offenbar bezüglich der Addition ein Isomorphismus. Man kann ihn auf die Multiplikation ausdehnen, indem man für die $f(u, v^{-1})$ ein Produkt in folgender Weise definiert:

Unter $f(u, v^{-1}) \times g(u, v^{-1})$ werde das Polynom in u, v^{-1} verstanden, das man durch formales Ausmultiplizieren und Voranstellen der Faktoren u erhält.

Die partiellen Ableitungen sollen für $f = \sum c_{ik} u^i v^{-k}$ durch

$$\frac{\partial f}{\partial u} = \sum c_{ik}\, i\, u^{i-1} v^{-k}, \qquad \frac{\partial f}{\partial v} = \sum c_{ik} (-k)\, u^i v^{-k-1}$$

definiert sein. (Dabei gilt offenbar die Produktregel, z. B.

$$\frac{\partial}{\partial u}(f \times g) = \frac{\partial f}{\partial u} \times g + f \times \frac{\partial g}{\partial u}\Big).$$

Das Produkt $f \cdot g$ soll dann sein:

$$f \cdot g = \sum_{\nu} \frac{1}{\nu!} \frac{\partial^\nu f}{\partial v^\nu} \times \frac{\partial^\nu g}{\partial u^\nu}.$$

Von diesem Produkt läßt sich beweisen, daß es distributiv und assoziativ ist. Dies wird später in noch größerer Allgemeinheit geschehen.

Bezüglich dieses Produktes für die Polynome $f(u, v^{-1})$ ist nun die Zuordnung

$$f(s, I) \leftrightarrow f(u, v^{-1})$$

ein Isomorphismus: Nach der getroffenen Produktdefinition erfüllen nämlich die Elemente u, v^{-1} die Vertauschungsrelation:

$$v^{-1} u = u v^{-1} - (v^{-1})^2.$$

Man rechnet also mit den Polynomen $f(u, v^{-1})$ ebenso wie mit den Polynomen $f(s, I)$.

Wir haben also folgenden Isomorphiesatz:

Der Ring $\mathsf{K}[s, I]$, *bestehend aus den Operatorpolynomen* $\sum c_{ik} s^i I^k$, *wird vermöge der Zuordnung*

$$\sum c_{ik} s^i I^k \to \sum c_{ik} u^i v^{-k}$$

isomorph auf den Polynomring $\mathsf{K}[u, v^{-1}]$ *bezogen, wenn für letzteren obige Produktdefinition zugrunde gelegt wird:*

$$\mathsf{K}[s, I] \cong \mathsf{K}[u, v^{-1}].$$

§ 2. Integraloperatoren und ihre Kerne

DEFINITION. *Ein Operator aus* $\mathsf{K}[s, I]$ *heiße Integraloperator, wenn er das von I freie Glied nicht enthält, also von der Form*

$$\varphi_0 I + \varphi_1 I^2 + \cdots + \varphi_{n-1} I^n$$

ist, wobei die φ_i *Polynome in s sind.*

Die Ausübung eines solchen Integraloperators K kann auf eine einzige Integration zurückgeführt werden:

$$K f = \int_0^s \left(\varphi_0(s) + \varphi_1(s) \frac{(s-t)}{1!} + \cdots + \varphi_{n-1}(s) \frac{(s-t)^{n-1}}{(n-1)!} \right) f(t)\, dt$$

$$= \int_0^s K(s,t)\, f(t)\, dt \quad \text{mit} \quad K(s,t) = \sum_{\nu=0}^{n-1} \varphi_\nu(s) \frac{(s-t)^\nu}{\nu!}.$$

DEFINITION. *Ist*

$$K = \varphi_0 I + \varphi_1 I^2 + \cdots + \varphi_{n-1} I^n,$$

so bezeichnen wir

$$K(s,t) = \varphi_0(s) + \varphi_1(s) \frac{(s-t)}{1!} + \cdots + \varphi_{n-1}(s) \frac{(s-t)^{n-1}}{(n-1)!}$$

als den zum Integraloperator gehörigen Kern.

Zwischen den Integraloperatoren und den zugehörigen Kernen besteht eine umkehrbar eindeutige Zuordnung

$$K \sim K(s,t)$$

und diese ist offenbar bezüglich der Addition ein Isomorphismus.

Bezüglich der Multiplikation gilt der

SATZ. *Aus*

$$K_i \sim K_i(s,t), \qquad i = 1, 2$$

folgt:

$$K_1 K_2 \sim \int_t^s K_1(s,\sigma)\, K_2(\sigma,t)\, d\sigma.$$

Dem Operatorprodukt $K_1 K_2$ entspricht also die „Volterrasche Komposition" von $K_1(s,t)$, $K_2(s,t)$.

Der Beweis ist aus der Theorie der Integralgleichungen bekannt. Der letzte Satz gibt Veranlassung zu der

DEFINITION. *Sind $K_1(s,t)$, $K_2(s,t)$ Polynome in s, t, so soll sein:*

$$K_1(s,t) * K_2(s,t) = \int_t^s K_1(s,\sigma)\, K_2(\sigma,t)\, d\sigma.$$

Bezüglich dieses Produkts ist dann die Zuordnung

$$K \to K(s,t)$$

ausführlich:

$$\varphi_0(s)\, I + \cdots + \varphi_{n-1}(s)\, I^n \to \varphi_0(s) + \cdots + \varphi_{n-1}(s) \frac{(s-t)^{n-1}}{(n-1)!}$$

ein Isomorphismus.

Andererseits können wir die Integraloperatoren K nach §1 auch zum Ring $K[u, v^{-1}]$ in Isomorphie bringen, und zwar mit der Zuordnung:

$$\varphi_0(s)\, I + \cdots + \varphi_{n-1}(s)\, I^n \to \varphi_0(u)\, v^{-1} + \cdots + \varphi_{n-1}(u)\, v^{-n}.$$

Hieraus folgt der

SATZ. *Ordnet man dem Kern*

$$K(s,t) = \varphi_0(s) + \cdots + \varphi_{n-1}(s)\,\frac{(s-t)^{n-1}}{(n-1)!}$$

das Element

$$\varphi_0(u)\,v^{-1} + \cdots + \varphi_{n-1}(u)\,v^{-n} \in \mathsf{K}[u, v^{-1}]$$

zu, so ergibt sich ein Isomorphismus. Bei diesem geht die Volterrasche Komposition $K_1(s,t) * K_2(s,t)$ *in das Produkt der Bildelemente über.*

Diese Zuordnung ist in spezieller Form bereits diejenige Transformation $\mathfrak{T}$, die die Rolle der Laplace-Transformation übernehmen soll. Für die Anwendung auf Differentialgleichungen sind noch folgende Verallgemeinerungen nötig:

1. Der Ring $\mathsf{K}[u, v^{-1}]$ ist in einen Schiefkörper einzubetten.

2. Während bisher die Kerne $K(s,t)$ Polynome in s, t waren, sollen es künftig komplexe Funktionen sein, die für

$$|s| < \delta, \qquad |t| < \delta$$

analytisch sind.

§ 3. Einbettung von $\mathsf{K}[u, v^{-1}]$ in einen Schiefkörper

Unter $\varphi(u)$ werde ein Ausdruck

$$c_0 u^n + c_1 u^{n+1} + c_2 u^{n+2} + \cdots \quad \text{mit} \quad c_0 \neq 0$$

verstanden, wobei n eine ganze Zahl $\gtreqless 0$ und die c_i komplexe Zahlen. Zur Menge dieser Elemente nehmen wir die Null hinzu. Zwei solche Elemente heißen gleich, wenn sie entweder beide gleich Null oder wenn sie in den Zahlen n, c_i übereinstimmen. Die Addition zweier solcher Elemente soll durch Addition entsprechender c_i geschehen. Das Produkt soll durch formales Ausmultiplizieren gebildet werden: Ist

$$\varphi(u) = c_0 u^n + c_1 u^{n+1} + \cdots$$
$$\bar{\varphi}(u) = \bar{c}_0 u^{\bar{n}} + \bar{c}_1 u^{\bar{n}+1} + \cdots,$$

so definieren wir:

$$\varphi(u)\,\bar{\varphi}(u) = c_0\bar{c}_0\,u^{n+\bar{n}} + (c_0\bar{c}_1 + c_1\bar{c}_0)\,u^{n+\bar{n}+1} + (c_0\bar{c}_2 + c_1\bar{c}_1 + c_2\bar{c}_0)\,u^{n+\bar{n}+2} + \cdots.$$

Mit diesen Verknüpfungen ergibt sich offenbar ein kommutativer Ring. Ferner: Zu jedem $\varphi(u) \neq 0$ existiert das Inverse φ^{-1}. Setzt man nämlich

$$\varphi^{-1} = c_0' u^{n'} + c_1' u^{n'+1} + \cdots,$$

so ergibt sich $n' = -n$ und das folgende Rekursionssystem für die c_i':

$$c_0 c_0' = 1$$
$$c_0 c_1' + c_1 c_0' = 0$$
$$\sum c_\nu c_{n-\nu}' = 0 \qquad (n \geqq 1).$$

Wegen $c_0 \neq 0$ läßt es sich eindeutig nach den c_i' auflösen. Die Menge der $\varphi(u)$ bilden also einen kommutativen Körper: $\mathsf{K}(u)$.

Für das Folgende definieren wir noch für $\varphi = \sum_0^\infty c_i\, u^{n+i}$ die Ableitung:

$$\varphi'(u) = \sum_0^\infty c_i (n+i)\, u^{n+i-1}.$$

Ausgehend von $\mathsf{K}(u)$ konstruieren wir nun eine neue Art von Elementen:

Es sei $\varphi_i(u)$ eine unendliche Folge $\in \mathsf{K}(u)$, $\varphi_0(u) \neq 0$ und m eine ganze Zahl $\gtreqless 0$. Wir bilden den Ausdruck

$$f(u, v) = \varphi_0(u)\, v^m + \varphi_1(u)\, v^{m-1} + \varphi_2(u)\, v^{m-2} + \cdots.$$

Es ist zu beachten, daß die Exponenten von v fallen, während die von u steigen. Zu der Menge dieser Ausdrücke nehmen wir die Null hinzu.

Wir definieren nun:

$f = \bar{f}$, wenn entweder beide gleich Null, oder wenn $m = \bar{m}$ und $\varphi_i(u) = \bar{\varphi}_i(u)$, $i = 0, 1, 2, \ldots$

$f_1 + f_2$ wird durch Addition entsprechender Koeffizienten $\varphi(u)$ erhalten.

Ferner soll ein nichtkommutatives Produkt definiert werden. Zu diesem Zweck führen wir folgende Hilfsbegriffe ein:

a) *Das Normalprodukt* $f(u, v) \times g(u, v)$ soll durch „normales" Ausmultiplizieren gebildet werden:

Ist

$$f = \varphi_0(u)\, v^m + \varphi_1(u)\, v^{m-1} + \cdots$$
$$g = \psi_0(u)\, v^n + \psi_1(u)\, v^{n-1},$$

so soll sein:

$$f \times g = \varphi_0 \psi_0\, v^{m+n} + (\varphi_0 \psi_1 + \varphi_1 \psi_0)\, v^{m+n-1} + (\varphi_0 \psi_2 + \varphi_1 \psi_1 + \varphi_2 \psi_0)\, v^{m+n-2} + \cdots.$$

b) *Partielle Ableitungen* sind rein formal zu bilden:

$$\frac{\partial f}{\partial u} = \varphi_0'(u)\, v^m + \varphi_1'(u)\, v^{m-1} + \cdots$$
$$\frac{\partial f}{\partial v} = m\, \varphi_0(u)\, v^{m-1} + (m-1)\, \varphi_1(u)\, v^{m-2} + \cdots.$$

c) Es sei $f_i(u, v)$ eine unendliche Folge:

$$f_1 = \varphi_0^{(1)} v^{n_1} + \varphi_1^{(1)} v^{n_1 - 1} + \cdots \qquad \varphi_0^{(1)} \neq 0$$
$$f_2 = \varphi_0^{(2)} v^{n_2} + \varphi_2^{(2)} v^{n_2 - 1} + \cdots \qquad \varphi_0^{(2)} \neq 0$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

bei der die n_i monoton fallen:

$$n_1 > n_2 > n_3 > \cdots.$$

Jede fest gewählte Potenz von v kommt dann höchstens in endlich vielen $f_i(u, v)$ vor und man kann daher in sinnvoller Weise *die unendliche Summe*

$$f_1 + f_2 + f_3 + \cdots$$

definieren durch die Vorschrift, daß entsprechende Koeffizienten zu addieren sind.

Die Definition des nichtkommutativen Produkts lautet nun:

$$f(u,v) \cdot g(u,v) = f \times g + \frac{1}{1!} \frac{\partial f}{\partial v} \times \frac{\partial g}{\partial u} + \frac{1}{2!} \frac{\partial^2 f}{\partial v^2} \times \frac{\partial^2 g}{\partial u^2} + \cdots .$$

Nach c) ist diese unendliche Summe sinnvoll, da wegen der partiellen Ableitungen $\partial^\nu f / \partial v^\nu$ die n_i monoton fallen.

BEWEIS DER KÖRPERGESETZE.

Im folgenden soll nachgewiesen werden, daß die Menge der $f(u, v)$ bezüglich der angegebenen Verknüpfungen einen nichtkommutativen Körper bilden.

Hierfür werden zunächst einige Hilfsmittel zusammengestellt.

DEFINITION. *Ist* $f(u,v) = \varphi_0(u) v^m + \varphi_1(u) v^{m-1} + \cdots$ *mit* $\varphi_0(u) \neq 0$, *so soll* m *der Grad von* f *heißen.*

DEFINITION. *Die Elemente* $f(u, v)$, $g(u, v)$ *heißen* p*-äquivalent, in Zeichen*

$$f \underset{p}{\sim} g,$$

wenn $f - g$ *entweder gleich Null oder höchstens vom Grad* $p - 1$.

Für jedes feste p ist diese Relation offenbar eine Äquivalenzrelation.

HILFSSATZ 1. *Ist* $f \underset{p}{\sim} g$ *für jedes beliebige — noch so kleine —* p *erfüllt, so ist* $f = g$ *und umgekehrt.* Beweis trivial.

HILFSSATZ 2. *Hat das Produkt* $f\,g$ *den Grad* M *und ist* k *eine ganze Zahl* $\geqq 0$, *so gilt*

$$f\,g \underset{M-k}{\sim} \sum_{\nu=0}^{k} \frac{1}{\nu!} \frac{\partial^\nu f}{\partial v^\nu} \times \frac{\partial^\nu g}{\partial u^\nu} .$$

BEWEIS. Nach Definition ist

$$f\,g = \sum_{\nu=0}^{\infty} \frac{1}{\nu!} \frac{\partial^\nu f}{\partial v^\nu} \times \frac{\partial^\nu g}{\partial u^\nu} ,$$

also

$$f\,g - \sum_{\nu=0}^{k} \frac{1}{\nu!} \frac{\partial^\nu f}{\partial v^\nu} \times \frac{\partial^\nu g}{\partial u^\nu} = \sum_{\nu=k+1}^{\infty} \frac{1}{\nu!} \frac{\partial^\nu f}{\partial v^\nu} \times \frac{\partial^\nu g}{\partial u^\nu} .$$

Die rechte Seite ist entweder gleich Null oder hat höchstens den Grad $M - k - 1$ q.e.d.

Nach diesen Vorbereitungen nun zum Beweis der Körpergesetze. Gleichheit und Summe haben offenbar die üblichen Eigenschaften.

Beweis des distributiven Gesetzes $f(g+h)=fg+fh$. Wir nehmen zunächst $f(g+h)\neq 0$ an, etwa vom Grad M, und wählen irgendeine ganze Zahl $k\geqq 0$.

Nach Hilfssatz 2 ist dann:

$$f(g+h)\underset{M-k}{\sim}\sum_{\nu=0}^{k}\frac{1}{\nu!}\,\frac{\partial^\nu f}{\partial v^\nu}\times\frac{\partial^\nu}{\partial u^\nu}(g+h)$$

$$=\sum_{\nu=0}^{k}\frac{1}{\nu!}\,\frac{\partial^\nu f}{\partial v^\nu}\times\frac{\partial^\nu g}{\partial u^\nu}+\sum_{\nu=0}^{k}\frac{1}{\nu!}\,\frac{\partial^\nu f}{\partial v^\nu}\times\frac{\partial^\nu h}{\partial u^\nu}$$

$$\underset{M-k}{\sim} fg+fh.$$

Da hier k jede ganze Zahl $\geqq 0$ sein kann, ist nach Hilfssatz 1

$$f(g+h)=fg+fh.$$

Der Fall $f(g+h)=0$ läßt sich wegen der Nullteilerfreiheit des Produkts ebenfalls leicht erledigen.

Ganz entsprechend kann man den Beweis für $(f+g)h=fh+gh$ führen.

Beweis des Assoziativgesetzes:

Wir nehmen dabei an, daß keiner der Faktoren f, g, h verschwindet. Sei M der Grad von $f(gh)$ und k wieder eine beliebig gewählte ganze Zahl $\geqq 0$. Wir geben zuerst ein Element A an mit

$$f(gh)\underset{M-k}{\sim}A,$$

dann ein Element A' mit

$$(fg)\,h\underset{M-k}{\sim}A'$$

und zeigen schließlich $A=A'$.

(a)
$$f(gh)\underset{M-k}{\sim}\sum_{\nu=0}^{k}\frac{1}{\nu!}\,\frac{\partial^\nu f}{\partial v^\nu}\times\frac{\partial^\nu (gh)}{\partial u^\nu}$$

$$=\sum_{\nu=0}^{k}\frac{1}{\nu!}\,\frac{\partial^\nu f}{\partial v^\nu}\times\frac{\partial^\nu}{\partial u^\nu}\left(\sum_{\lambda=0}^{\infty}\frac{1}{\lambda!}\,\frac{\partial^\lambda g}{\partial v^\lambda}\times\frac{\partial^\lambda h}{\partial u^\lambda}\right)$$

$$\underset{M-k}{\sim}\sum_{\nu=0}^{k}\frac{1}{\nu!}\,\frac{\partial^\nu f}{\partial v^\nu}\times\frac{\partial^\nu}{\partial u^\nu}\sum_{\lambda=0}^{k-\nu}\frac{1}{\lambda!}\,\frac{\partial^\lambda g}{\partial v^\lambda}\times\frac{\partial^\lambda h}{\partial u^\lambda}.$$

Wegen der partiellen Ableitungen nach v haben nämlich die fortgelassenen Glieder höchstens den Grad $M-k-1$. Es ist also:

$$f(gh)\underset{M-k}{\sim}\sum_{\nu=0}^{k}\frac{1}{\nu!}\,\frac{\partial^\nu f}{\partial v^\nu}\times\sum_{\lambda=0}^{k-\nu}\frac{1}{\lambda!}\sum_{\sigma=0}^{\nu}\binom{\nu}{\sigma}\frac{\partial^{\lambda+\nu-\sigma}g}{\partial v^\lambda\,\partial u^{\nu-\sigma}}\times\frac{\partial^{\lambda+\sigma}h}{\partial u^{\lambda+\sigma}}$$

$$=\sum\frac{1}{\nu!}\,\frac{1}{\lambda!}\binom{\nu}{\sigma}\frac{\partial^\nu f}{\partial v^\nu}\times\frac{\partial^{\lambda+\nu-\sigma}g}{\partial v^\lambda\,\partial u^{\nu-\sigma}}\times\frac{\partial^{\lambda+\sigma}h}{\partial u^{\lambda+\sigma}}=A.$$

Zu summieren ist über die Tripel ν, λ, σ mit

1. $\nu, \lambda, \sigma \geqq 0$. 2. $\nu \leqq k$. 3. $\nu + \lambda \leqq k$. 4. $\sigma \leqq \nu$.

(b)
$$(f\,g)\,h \underset{M-k}{\sim} \sum_{\nu=0}^{k} \frac{1}{\nu!} \frac{\partial^\nu (f\,g)}{\partial v^\nu} \times \frac{\partial^\nu h}{\partial u^\nu}$$

$$\underset{M-k}{\sim} \sum_{\nu=0}^{k} \frac{1}{\nu!} \frac{\partial^\nu}{\partial v^\nu} \left(\sum_{\lambda=0}^{k-\nu} \frac{1}{\lambda!} \frac{\partial^\lambda f}{\partial v^\lambda} \times \frac{\partial^\lambda g}{\partial u^\lambda} \right) \times \frac{\partial^\nu h}{\partial u^\nu}$$

$$= \sum_{\nu=0}^{k} \frac{1}{\nu!} \sum_{\lambda=0}^{k-\nu} \frac{1}{\lambda!} \sum_{\sigma=0}^{\nu} \binom{\nu}{\sigma} \frac{\partial^{\lambda+\sigma} f}{\partial v^{\lambda+\sigma}} \times \frac{\partial^{\nu+\lambda-\sigma} g}{\partial u^\lambda \partial v^{\nu-\sigma}} \times \frac{\partial^\nu h}{\partial u^\nu}$$

$$= \sum \frac{1}{\nu!} \frac{1}{\lambda!} \binom{\nu}{\sigma} \frac{\partial^{\lambda+\sigma} f}{\partial v^{\lambda+\sigma}} \times \frac{\partial^{\lambda+\nu-\sigma} g}{\partial u^\lambda \partial v^{\nu-\sigma}} \times \frac{\partial^\nu h}{\partial u^\nu} = A'.$$

c) Um zu zeigen, daß $A = A'$, führen wir in A' neue Summationsindizes ein:

$$\bar{\nu} = \lambda + \sigma, \qquad \bar{\lambda} = \nu - \sigma, \qquad \bar{\sigma} = \sigma.$$

Aus A' wird dann:

$$\sum \frac{1}{(\bar{\lambda}+\bar{\sigma})!} \frac{1}{(\bar{\nu}-\bar{\sigma})!} \binom{\bar{\lambda}+\bar{\sigma}}{\bar{\sigma}} \frac{\partial^{\bar{\nu}} f}{\partial v^{\bar{\nu}}} \times \frac{\partial^{\bar{\nu}+\bar{\lambda}-\bar{\sigma}} g}{\partial u^{\bar{\nu}-\bar{\sigma}} \partial v^{\bar{\lambda}}} \times \frac{\partial^{\bar{\lambda}+\bar{\sigma}} h}{\partial u^{\bar{\lambda}+\bar{\sigma}}}$$

$$= \sum \frac{1}{\bar{\nu}!} \frac{1}{\bar{\lambda}!} \binom{\bar{\nu}}{\bar{\sigma}} \frac{\partial^{\bar{\nu}} f}{\partial v^{\bar{\nu}}} \times \frac{\partial^{\bar{\nu}+\bar{\lambda}-\bar{\sigma}} g}{\partial u^{\bar{\nu}-\bar{\sigma}} \partial v^{\bar{\lambda}}} \times \frac{\partial^{\bar{\lambda}+\bar{\sigma}} h}{\partial u^{\bar{\lambda}+\bar{\sigma}}}.$$

Die $\bar{\nu}, \bar{\lambda}, \bar{\sigma}$ durchlaufen dabei ebenfalls sämtliche Tripel mit den Eigenschaften 1. bis 4. wie unmittelbar an den Substitutionsgleichungen zu sehen. Also ist $A = A'$ und somit

$$f\,(g\,h) \underset{M-k}{\sim} (f\,g)\,h.$$

Da k hier wieder jede ganze Zahl $\geqq 0$ sein kann, besagt dies

$$f\,(g\,h) = (f\,g)\,h, \quad \text{q.e.d.}$$

EXISTENZ DES INVERSEN:

Die Menge der $f(u, v)$ ist als Körper erwiesen, wenn noch gezeigt ist, daß es zu jedem $f(u, v) \neq 0$ ein Element $g(u, v)$ gibt mit $fg = 1$.

Zum Beweise nehmen wir an:

$$f = \varphi_0(u)\, v^m + \varphi_1(u)\, v^{m-1} + \cdots, \qquad \varphi_0 \neq 0$$

und machen für g den Ansatz

$$g = \psi_0(u)\, v^n + \psi_1(u)\, v^{n-1} + \cdots.$$

Hierin ist notwendig $n = -m$ zu setzen, da der Grad des Produkts gleich $m+n$ ist, andererseits gleich 0 werden soll. Im übrigen erhalten wir durch Ausmultiplizieren gemäß der Produktdefinition die folgenden Bestimmungsgleichungen für die Unbekannten ψ_i:

$$\varphi_0 \psi_0 = 1$$

$$(\varphi_0 \psi_1 + \varphi_1 \psi_0) + m\, \varphi_0 \psi_0' = 0$$

$$(\varphi_0 \psi_2 + \varphi_1 \psi_1 + \varphi_2 \psi_0) + \frac{1}{1!} \left(m\, \varphi_0 \psi_1' + (m-1)\, \varphi_1 \psi_0' \right) + \frac{m(m-1)}{2!} \varphi_0 \psi_0'' = 0$$

. . . . . .

Hieraus lassen sich die ψ_i rekursiv berechnen ($\varphi_0 \neq 0$). Aus der 0-ten Gleichung erhält man $\psi_0 = \frac{1}{\varphi_0}$ und allgemein aus der n-ten ψ_n, denn diese enthält ψ_n mit dem Faktor φ_0 und sonst von den ψ_i nur mehr die schon festgelegten $\psi_0, \ldots, \psi_{n-1}$. Es ist somit auch die Existenz des Inversen gesichert.

Die Menge der $f(u, v)$ bilden demnach einen Körper $\mathsf{SK} = \mathsf{K}(u, v)$. SK *enthält einen zum Ring* $\mathsf{K}[u, v^{-1}]$ *isomorphen Unterring, und zwar wird dieser gebildet von den speziellen Elementen*

$$\varphi_0(u) + \varphi_1(u)\, v^{-1} + \varphi_2(u)\, v^{-2} + \cdots + \varphi_n(u)\, v^{-n},$$

deren Koeffizienten $\varphi_i(u)$ Polynome in u sind. Es ist damit die Einbettung des Rings $\mathsf{K}[u, v^{-1}]$ *in einen Körper vollzogen.*

§ 4. Die Transformation $\mathfrak{T}$

In §2 wurde der von den Polynomkernen $K(s, t)$ gebildete Ring isomorph in den Ring $\mathsf{K}[u, v^{-1}]$ abgebildet vermöge

$$K(s,t) = \varphi_0(s) + \varphi_1(s)\frac{(s-t)}{1!} + \cdots + \varphi_n(s)\frac{(s-t)^n}{n!} \to \varphi_0(u)\, v^{-1} + \cdots + \varphi_n(u)\, v^{-n-1}.$$

Der Volterraschen Komposition der Urelemente entspricht dabei das Produkt der Bildelemente.

Diese Abbildung soll nun erweitert werden auf die Klasse C_δ der für

$$|s| < \delta, \qquad |t| < \delta$$

analytischen Funktionen $F(s, t)$. Die Bilder werden dabei dem Schiefkörper SK entnommen.

Definition von $\mathfrak{T}$. *Man entwickle die Funktion $F(s, t) \in C_\delta$ in bezug auf t nach Potenzen von $(s - t)$:*

$$F(s,t) = \varphi_0(s) + \varphi_1(s)\frac{(s-t)}{1!} + \varphi_2(s)\frac{(s-t)^2}{2!} + \cdots.$$

Diese Entwicklung ist für jedes s mit $|s| < \delta$ eindeutig, und zwar ist nach Taylor:

$$\varphi_\nu(s) = (-1)^\nu \left.\frac{\partial^\nu F}{\partial t^\nu}\right|_{t=s}$$

An Hand dieser Entwicklung definieren wir die Zuordnung $\mathfrak{T}$:

$$\mathfrak{T}\{F\} = \varphi_0(u)\, v^{-1} + \varphi_1(u)\, v^{-2} + \cdots$$

Das Bild liegt dabei in SK.

$\mathfrak{T}$ ist offenbar bezüglich der Addition ein Isomorphismus. Um auch bezüglich der Multiplikationen einen Isomorphismus zu erreichen, geben wir die

Definition. *Sind $F, G \in C_\delta$, so soll sein*

$$F * G = \int_t^s F(s, \sigma)\, G(\sigma, t)\, d\sigma, \qquad |s| < \delta, \qquad |t| < \delta.$$

Dieses Produkt gehört offenbar ebenfalls zu C_δ. Es gilt nun der

SATZ:

$$\mathfrak{L}\{F * G\} = \mathfrak{L}\{F\}\,\mathfrak{L}\{G\}.$$

Es geht also die Volterrasche Komposition der analytischen Funktionen $F(s,t)$, $G(s,t)$ in das Produkt ihrer Bilder über.

BEWEIS. Wir werden einerseits für die Entwicklung

$$\int_t^s F(s,\sigma)\,G(\sigma,t)\,d\sigma = \varrho_0(s)\,\frac{(s-t)}{1!} + \varrho_1(s)\,\frac{(s-t)^2}{2!} + \cdots$$

die $\varrho_\nu(s)$ berechnen, andererseits für das Produkt der Bilder

$$\mathfrak{L}\{F\}\,\mathfrak{L}\{G\} = \chi_0(u)\,v^{-2} + \chi_1(u)\,v^{-3} + \cdots$$

die χ_ν, wobei sich $\varrho_\nu = \chi_\nu$ ergeben wird.

a) Setzt man

$$H(s,t) = \int_t^s F(s,\sigma)\,G(\sigma,t)\,d\sigma,$$

so ist

$$\varrho_\nu = (-1)^{\nu+1}\,\frac{\partial^{\nu+1}H}{\partial t^{\nu+1}}\bigg|_{t=s}$$

Explizit:

$$-\frac{\partial H}{\partial t} = F(s,t)\,G(t,t) + \int_s^t F(s,\sigma)\,\frac{\partial G(\sigma,t)}{\partial t}\,d\sigma$$

$$-\frac{\partial^2 H}{\partial t^2} = \frac{\partial}{\partial t}\left(F(s,t)\,G(t,t)\right) + F(s,t)\,\frac{\partial G(\sigma,t)}{\partial t}\bigg|_{\sigma=t} + \int_s^t F(s,\sigma)\,\frac{\partial^2 G(\sigma,t)}{\partial t^2}\,d\sigma$$

. . . . . .

$$-\frac{\partial^{\nu+1}H}{\partial t^{\nu+1}} = \frac{\partial^\nu}{\partial t^\nu}\left(F(s,t)\,G(t,t)\right) + \frac{\partial^{\nu-1}}{\partial t^{\nu-1}}\left(F(s,t)\,\frac{\partial G(\sigma,t)}{\partial t}\bigg|_{\sigma=t}\right) +$$

$$+ \cdots + F(s,t)\,\frac{\partial^\nu G(\sigma,t)}{\partial t^\nu}\bigg|_{\sigma=t} + \int_s^t F(s,\sigma)\,\frac{\partial^{\nu+1}G(\sigma,t)}{\partial t^{\nu+1}}\,d\sigma.$$

Hierbei ist:

$$\frac{\partial^\nu F}{\partial t^\nu}(s,t) = (-1)^\nu\left(\varphi_\nu(s) + \varphi_{\nu+1}\,\frac{(s-t)}{1!} + \cdots\right), \qquad \frac{\partial^\nu G(\sigma,t)}{\partial t^\nu}\bigg|_{\sigma=t} = (-1)^\nu\,\psi_\nu(t).$$

Dies geht für $t=s$ über in:

$$(-1)^\nu\,\varphi_\nu(s) \quad \text{bzw.} \quad (-1)^\nu\,\psi_\nu(s).$$

Setzt man symbolisch

$$(-1)^\nu\,\varphi_\nu(s) = \varphi^{(\nu)},$$

so erhält man schließlich:

$$(-1)^\nu\,\varrho_\nu(s) = -\frac{\partial^{\nu+1}H}{\partial t^{\nu+1}}\bigg|_{t=s} = (\varphi_0\psi_0)^{(\nu)} - (\varphi_0\psi_1)^{(\nu-1)} + \cdots + (-1)^\nu\,\varphi_0\psi_\nu.$$

b) Wir berechnen die ψ_ν gemäß der Produktdefinition

$$fg = \sum_{\nu=0}^{\infty} \frac{1}{\nu!} \frac{\partial^\nu f}{\partial v^\nu} \times \frac{\partial^\nu g}{\partial u^\nu}:$$

$$\chi_0 = \varphi_0 \psi_0$$

$$(-1)\chi_1 = (-\varphi_1\psi_0 + \varphi_0\psi_0') + \varphi_0\psi_1$$

$$(-1)^2\chi_2 = \left(\binom{2}{0}\varphi_2\psi_0 - \binom{2}{1}\varphi_1\psi_0' + \binom{2}{2}\varphi_0\psi_0''\right) - (-\varphi_1\psi_1 + \varphi_0\psi_1') + \varphi_0\psi_2.$$

Es ist hierbei nach dem Index von ψ geordnet worden. Setzen wir symbolisch

$$(-1)^\nu \varphi_\nu = \varphi^{(\nu)},$$

so können wir einfacher schreiben:

$$(-1)^2\chi_2 = (\varphi_0\psi_0)'' - (\varphi_0\psi_1)' + (\varphi_0\psi_2)^{(0)}.$$

Allgemein liefert die Produktdefinition:

$$(-1)^\nu\chi_\nu = (\varphi_0\psi_0)^\nu - (\varphi_0\psi_1)^{\nu-1} + \cdots + (-1)^\nu\varphi_0\psi_\nu.$$

Es ist also tatsächlich

$$\varphi_\nu = \chi_\nu$$

und somit

$$\mathfrak{T}\{F * G\} = \mathfrak{T}\{F\}\mathfrak{T}\{G\}, \quad \text{q.e.d.}$$

§ 5. Umkehrung der Transformation $\mathfrak{T}$

$\mathfrak{T}$ ordnet jeder analytischen Funktion aus C_δ in eindeutiger Weise ein Element aus dem Schiefkörper SK zu. Im folgenden taucht die Frage auf, ob es umgekehrt zu einem vorgegebenen Element $g \in \mathsf{SK}$ eine Funktion $G(s,t)$ mit $\mathfrak{T}\{G\} = g$ gibt. Es existiert offenbar nicht zu jedem $g \in \mathsf{SK}$ eine solche Urfunktion G, denn hierfür ist zunächst notwendig, daß g die Form $\varphi_0 v^{-1} + \varphi_1 v^{-2} + \cdots$ hat. Jedoch sind auch unter den Elementen dieser Art noch solche, die keine Urfunktion G besitzen, wie z.B. $g = \sum_{\nu=0}^{\infty} (\nu!)^2 v^{-\nu-1}$.

Definition. *Ein Element g aus SK heiße Funktionselement, wenn es dazu eine in einer gewissen Umgebung $|s| < \delta$, $|t| < \delta$ analytische Funktion $G(s,t)$ gibt mit*

$$\mathfrak{T}\{G\} = g.$$

Mit gebräuchlichen Mitteln beweist man den

Satz 5.1. *Ist $g \in \mathsf{SK}$ Funktionselement und sind $G_1(s,t)$, $G_2(s,t)$ beide Urfunktionen von g, etwa mit dem gemeinsamen Regularitätsbereich $|s| < \delta$, $|t| < \delta$, so sind sie dort identisch.*

In diesem Sinne ist also die Urfunktion eines Funktionselements eindeutig bestimmt.

Das zu einem Funktionselement g gehörige Urbild kann man aus g direkt bestimmen nach dem

SATZ 5.2. *Ist*

$$g = \varphi_0(u)\, v^{-1} + \varphi_1(u)\, v^{-2} + \cdots$$

Funktionselement und ist die zugehörige Urfunktion $G(s,t)$ *für* $|s|<\delta$, $|t|<\delta$ *analytisch, so gilt für* $|s| < \frac{\delta}{3}$, $|t| < \frac{\delta}{3}$

$$G(s,t) = \sum_{\nu=0}^{\infty} \varphi_\nu(s)\, \frac{(s-t)^\nu}{\nu!}.$$

BEWEIS. Bei $G(s,t)$ werde die Variable s fixiert, und zwar mit $|s| < \frac{\delta}{3}$ und $G(s,t)$ in bezug auf t nach $(s-t)$ entwickelt:

$$G(s,t) = \sum_{\nu=0}^{\infty} \overline{\varphi}_\nu(s)\, \frac{(s-t)^\nu}{\nu!}.$$

Diese Reihe konvergiert für $|s-t| < \frac{2\delta}{3}$, also bei unserer Wahl von s für $|t| < \frac{\delta}{3}$. Da $\mathfrak{T}\{G\} = g$ sein soll, ist $\overline{\varphi}_\nu = \varphi_\nu$. Für $|s| < \frac{\delta}{3}$, $|t| < \frac{\delta}{3}$ konvergiert also $\sum_{\nu=0}^{\infty} \varphi_\nu(s)\, \frac{(s-t)^\nu}{\nu!}$ gegen $G(s,t)$, q.e.d.

Ferner brauchen wir im folgenden den

SATZ 5.3. *Sind* f, g *Funktionselemente, etwa mit den Urfunktionen* F, G, *die in* $|s|<\delta$, $|t|<\delta$ *analytisch sein mögen, so sind auch* $f \pm g$, fg *Funktionselemente, und zwar gehören dazu bzw. die Funktionen* $F \pm G$, $\int_t^s F(s,\sigma)\, G(\sigma,t)\, d\sigma$.

Beweis trivial.

Das Inverse g^{-1} eines von Null verschiedenen Funktionselements g ist nicht wieder Funktionselement, da g^{-1} einen positiven Grad hat. Dagegen kann es sein, daß g^{-1} Funktionselement ist, wenn es g selbst nicht ist. Für die Auflösung der linearen Differentialgleichung ist grundlegend der

SATZ 5.4. *Ist bei dem Element aus*

$$h = v^n + a_1(u)\, v^{n-1} + \cdots + a_n(u)$$

n *eine natürliche Zahl und sind die* $a_i(u)$ *so beschaffen, daß daraus durch die Ersetzung* $u \to s$ *Funktionen* $a_i(s)$ *entstehen, die sämtlich in einer Umgebung* $|s|<\delta$ *analytisch sind, so ist*

$$h^{-1} = v^{-n} + \psi_1\, v^{-n-1} + \cdots$$

Funktionselement.

Dem Beweis dient der

HILFSSATZ. *Hat* h *die Bedeutung des letzten Satzes, so gibt es eine „Zerlegung in Linearfaktoren"*

$$h = \big(v + \varphi_1(u)\big)\big(v + \varphi_2(u)\big) \ldots \big(v + \varphi_n(u)\big),$$

deren $\varphi_i(u)$ durch die Ersetzung $u \to s$ in Funktionen übergehen, die in einer gewissen Umgebung $|s| < \delta'$ analytisch sind.

BEWEIS. Neben h betrachte man den „Differentialoperator"

$$D^n + a_1(s)\, D^{n-1} + \cdots + a_n(s).$$

Dieser gestattet bekanntlich eine Zerlegung

$$\big(D + \varphi_1(s)\big) \ldots \big(D + \varphi_n(s)\big)^{2)} \tag{I}$$

mit Funktionen $\varphi_i(s)$, die in einer gewissen Umgebung $|s| < \delta$ analytisch sind. Diese Zerlegung kann man nun auf h übertragen. Ordnet man nämlich allgemein dem Differentialoperator

$$a_0(s)\, D^n + a_1(s)\, D^{n-1} + \cdots + a_n(s)$$

mit um 0 analytischen $a_i(s)$ das Element

$$a_0(u)\, v^n + a_1(u)\, v^{n-1} + \cdots + a_n(u)$$

zu, so ist dies offenbar ein Isomorphismus, weil einerseits für D die Vertauschungsrelation

$$D f(s) = f(s)\, D + f'(s)$$

gilt, andererseits für v die Vertauschungsrelation

$$v f(u) = f(u)\, v + f'(u),$$

was aus der Produktdefinition für SK folgt.

Nimmt man also in der Zerlegung (I) die Ersetzungen $s \to u$, $D \to v$ vor, so erhält man die behauptete Zerlegung von h.

Nun zum Beweis des letzten Satzes:

a) Sei $n = 1$, also

$$h = v + \varphi(u)$$

und

$$h^{-1} = v^{-1} + \chi_1 v^{-2} + \chi_2 v^{-3} + \cdots$$

$\varphi(s)$ sei für $|s| < \delta$ analytisch.

Wir zeigen, daß h^{-1} die Urfunktion

$$G(s,t) = e^{-\Phi(s)+\Phi(t)} \quad \text{mit} \quad \Phi(s) = \int_0^s \varphi(\sigma)\, d\sigma, \qquad |s| < \delta$$

besitzt.

Für $|s| < \delta$ und genügend benachbarte t gilt

$$G(s,t) = \sum_{\nu=0}^{\infty} \psi_\nu(s)\, \frac{(s-t)^\nu}{\nu!},$$

wobei

$$\psi_\nu(s) = (-1)^\nu \frac{\partial^\nu G}{\partial t^\nu}\bigg|_{t=s} = (-1)^\nu e^{-\Phi(s)} \frac{d^\nu e^{\Phi(s)}}{d s^\nu}.$$

---

2) Siehe etwa: SCHLESINGER, Handbuch der Theorie der linearen Differentialgleichungen.

Durch Differentiation nach s folgt die Rekursionsformel:

$$\psi_{\nu+1} + \psi_\nu \varphi + \psi_\nu' = 0.$$

Andererseits sind die χ_ν festgelegt durch

$$(v + \varphi)(v^{-1} + \chi_1 v^{-2} + \chi_2 v^{-3} + \cdots) = 1.$$

Explizit:

$$\chi_0 + (\chi_1 + \chi_0 \varphi + \chi_0') v^{-1} + \cdots + (\chi_{\nu+1} + \chi_\nu \varphi + \chi_\nu') v^{\nu+1} + \cdots = 1.$$

Hieraus folgt

$$\chi_0 = 1$$

$$\cdots\cdots\cdots$$

$$\chi_{\nu+1} + \chi_\nu \varphi + \chi_\nu' = 0.$$

Die χ_ν genügen also derselben Rekursionsformei wie die ψ_ν. Außerdem ist $\psi_0 = \chi_0 = 1$. Somit ist

$$\psi_\nu = \chi_\nu \quad \text{für} \quad \nu = 0, 1, 2, \ldots$$

und das besagt: $G(s, t)$ hat h^{-1} als Bild.

b) Im Falle $n > 1$ zerlegen wir:

$$h = (v + \varphi_1(u))(v + \varphi_2(u)) \ldots (v + \varphi_n(u)),$$

was nach dem Hilfssatz möglich ist. Die den $\varphi_i(u)$ entsprechenden $\varphi_i(s)$ seien dabei sämtlich für $|s| < \delta'$ analytisch. Nach a) gibt es zu jedem Faktor $v + \varphi_i(u)$ eine Urfunktion $G_i(s, t)$. Das Produkt hat daher nach §4 die Urfunktion

$$G_1 * G_2 * \ldots * G_n,$$

wobei dieses Produkt als Volterrasche Komposition zu lesen ist.

§ 6. Hilfssätze über die Transformation $\mathfrak{T}$

Die folgenden Hilfssätze werden für die Behandlung der linearen Differentialgleichung benötigt.

HILFSSATZ 6.1. *Ist für die um* 0 *analytische Funktion* $G(s, t)$

$$\mathfrak{T}\{G(s, t)\} = g,$$

so gilt:

$$\mathfrak{T}\left\{-\frac{\partial G}{\partial t}\right\} = g v - G(u, u),$$

wobei $G(u, u)$ *dasjenige Körperelement aus* SK *ist, das aus* $G(s, s)$ *durch die Ersetzung* $s \to u$ *hervorgeht.*

BEWEIS. Sei

$$G(s, t) = \psi_0(s) + \psi_1(s) \frac{(s-t)}{1!} + \psi_2(s) \frac{(s-t)^2}{2!} + \cdots$$

Dann ist

$$-\frac{\partial G}{\partial t} = \psi_1(s) + \psi_2 \frac{(s-t)^{1''}}{1!} + \cdots$$

und

$$\mathfrak{T}\{G\} = \psi_0(u)\, v^{-1} + \psi_1(u)\, v^{-2} + \cdots$$

$$\mathfrak{T}\left\{-\frac{\partial G}{\partial t}\right\} = \psi_1(u)\, v^{-1} + \psi_2(u)\, v^{-2} + \cdots$$

Zwischen den Bildern besteht offenbar die Beziehung

$$\mathfrak{T}\left\{-\frac{\partial G}{\partial t}\right\} = \mathfrak{T}\{G\}\, v - \psi_0(u),$$

die mit der Behauptung gleichbedeutend ist.

Wiederholte Anwendung ergibt:

$$(-1)^\nu \mathfrak{T}\left\{\frac{\partial^\nu G}{\partial t^\nu}\right\} = \mathfrak{T}\{G\}\, v^\nu - \psi_0(u)\, v^{\nu-1} - \cdots - \psi_{\nu-1}(u).$$

Folgerungen. a) Es sei $G(s,t)$ nur von $(s-t)$ abhängig:

$$G(s,t) = Y(s-t) = \sum_{i=0}^{\infty} c_i (s-t)^i.$$

$Y'(s-t)$ sei die Ableitung nach $(s-t)$. Dann gilt nach 6.1:

$$\mathfrak{T}\{Y'(s-t)\} = \mathfrak{T}\{Y\}\, v - Y(0)$$

und allgemein:

$$\mathfrak{T}\{Y^{(\nu)}(s-t)\} = \mathfrak{T}\{Y\}\, v^\nu - Y(0)\, v^{\nu-1} - \cdots - Y^{\nu-1}(0).$$

b) Hat $G(s,t)$ die Gestalt

$$\psi_k(s)\,\frac{(s-t)^k}{k!} + \psi_{k+1}(s)\,\frac{(s-t)^{k+1}}{(k+1)!} + \cdots,$$

wobei k eine natürliche Zahl, ist also $\psi_0 = \cdots \psi_{k-1} = 0$, so hat man nach 6.1 für $\nu = 1, \ldots, k$ die vereinfachte Formel:

$$(-1)^\nu \mathfrak{T}\left\{\frac{\partial^\nu G}{\partial t^\nu}\right\} = \mathfrak{T}\{G\}\, v^\nu.$$

Umgekehrt kann man schließen: Ist g ein Funktionselement von der Form

$$\psi_k(u)\, v^{-k-1} + \psi_{k+1}(u)\, v^{-k-2} + \cdots$$

mit der zugehörigen Urfunktion

$$G(s,t) = \psi_k(s)\,\frac{(s-t)^k}{k!} + \psi_{k+1}(s)\,\frac{(s-t)^{k+1}}{(k+1)!} + \cdots$$

und ist ν eine natürliche Zahl $\leqq k$, so ist auch $g\, v^\nu$ ein Funktionselement, nämlich nach Obigem das Abbild von

$$(-1)^\nu \frac{\partial^\nu G}{\partial t^\nu}.$$

Hilfssatz 6.2. *Sind $a(s)$ und $G(s,t)$ um 0 analytisch, so gilt*

$$\mathfrak{T}\{a(s)\, G(s,t)\} = a(u)\, \mathfrak{T}\{G(s,t)\}.$$

BEWEIS. Ist

$$G(s,t)=\sum_{\nu=0}^{\infty}\psi_\nu(s)\,\frac{(s-t)^\nu}{\nu!},$$

so gilt:

$$\mathfrak{T}\{a(s)\,G(s,t)\}=\mathfrak{T}\left\{\sum_{\nu=0}^{\infty}a(s)\,\psi_\nu(s)\,\frac{(s-t)^\nu}{\nu!}\right\}$$

$$=\sum_{\nu=0}^{\infty}a(u)\,\psi_\nu(u)\,v^{-\nu-1}=a(u)\sum_{\nu=0}^{\infty}\psi_\nu(u)\,v^{-\nu-1}$$

$$=a(u)\,\mathfrak{T}\{G(s,t)\}\quad\text{q.e.d.}$$

Ein ähnliches Transformationsgesetz gilt, wenn $a(t)$ eine nur von t abhängige Funktion ist.

HILFSSATZ 6.3. *Sind $G(s,t)$, $a(t)$ analytisch um 0 und hängt $a(t)$ nur von t ab, so gilt:*

$$\mathfrak{T}\{G(s,t)\,a(t)\}=\mathfrak{T}\{G(s,t)\}\cdot a(u).$$

BEWEIS. Wir nehmen den Sonderfall $G=1$ vorweg. Ist $a(t)$ für $|t|<\delta$ analytisch, so gilt für $|s|<\delta$ und benachbarte t:

$$a(t)=a(s)-a'(s)\,\frac{(s-t)}{1!}+a''(s)\,\frac{(s-t)^2}{2!}\mp\cdots.$$

Nach Definition von $\mathfrak{T}$ ist:

$$\mathfrak{T}\{a(t)\}=a(u)\,v^{-1}-a'(u)\,v^{-2}+a''(u)\,v^{-3}\mp\cdots.$$

Dieses Element ist aber identisch mit $v^{-1}\,a(u)$, wie die Produktdefinition für SK ergibt, d.h.

$$\mathfrak{T}\{a(t)\}=v^{-1}\,a(u).$$

Im allgemeinen Fall $G\neq 1$ gehen wir von der Identität

$$G(s,t)\,a(t)=\int_t^s\left(-\frac{\partial G(s,\sigma)}{\partial t}\right)a(t)\,d\sigma+G(s,s)\,a(t)$$

aus. Nach §4 und Hilfssatz 6.2 ergibt die Anwendung von $\mathfrak{T}$:

$$\mathfrak{T}\{G(s,t)\,a(t)\}=\mathfrak{T}\left\{-\frac{\partial G}{\partial t}\right\}\mathfrak{T}\{a(t)\}+G(u,u)\,\mathfrak{T}\{a(t)\};$$

nach Satz 6.1

$$=\left(\mathfrak{T}\{G\}\cdot v-G(u,u)\right)\mathfrak{T}\{a(t)\}+G(u,u)\,\mathfrak{T}\{a(t)\},$$

nach obigem Teilergebnis

$$=\mathfrak{T}\{G\}\,v\cdot v^{-1}\,a(u).$$

Also:

$$\mathfrak{T}\{G(s,t)\,a(t)\}=\mathfrak{T}\{G\}\,a(u),\quad\text{q.e.d.}$$

Schließlich sei noch ein Transformationsgesetz erwähnt, das einen zweiten einfachen Beweis für den grundlegenden Satz 5.4 ermöglicht.

Es gilt

$$\mathfrak{T}\left\{\frac{\partial G(s,t)}{\partial s}\right\} = v\,\mathfrak{T}\{G(s,t)\} - G(u,u).$$

BEWEIS. Es sei

$$G(s,t) = \sum_{\nu=0}^{\infty} \varrho_\nu(s)\,\frac{(s-t)^\nu}{\nu!},$$

also

$$\frac{\partial G}{\partial s} = \sum_{\nu=1}^{\infty} \varrho_\nu(s)\,\frac{(s-t)^{\nu-1}}{(\nu-1)!} + \sum_{\nu=0}^{\infty} \varrho_\nu'(s)\,\frac{(s-t)^\nu}{\nu!}.$$

Nach Definition von $\mathfrak{T}$ ist dann:

$$\begin{aligned}\mathfrak{T}\left\{\frac{\partial G}{\partial s}\right\} &= \sum_{\nu=1}^{\infty} \varrho_\nu(u)\,v^{-\nu} + \sum_{\nu=0}^{\infty} \varrho_\nu'(u)\,v^{-\nu-1}\\ &= \left(\sum_{\nu=0}^{\infty} \varrho_\nu(u)\,v^{-\nu-1}\right)v - \varrho_0(u) + \sum_{\nu=0}^{\infty} \varrho_\nu'(u)\,v^{-\nu-1}\\ &\doteq \mathfrak{T}\{G\}\,v + \frac{\partial}{\partial u}\,\mathfrak{T}\{G\} - G(u,u).\end{aligned}$$

Die beiden ersten Summanden ergeben nach der Produktdefinition für SK gerade $v\,\mathfrak{T}\{G\}$, womit das Transformationsgesetz bewiesen ist.

Es habe nunmehr h wieder dieselbe Bedeutung wie bei Satz 5.4:

$$h = v^n + a_1(u)\,v^{n-1} + \cdots + a_n(u),$$

wobei die Funktionen $a_i(s)$ der komplexen Veränderlichen s sämtlich um 0 analytisch sind.

Wir beweisen noch einmal, daß h^{-1} Funktionselement ist.

Wir gehen dabei von folgender bekannten Tatsache aus: Es existiert eine um 0 analytische Funktion $G(s,t)$ mit den Eigenschaften:

1. $G(s,t)$ ist in bezug auf s Lösung der homogenen Differentialgleichung:

$$\frac{\partial^n G(s,t)}{\partial s^n} + a_1(s)\,\frac{\partial^{n-1} G(s,t)}{\partial s^{n-1}} + \cdots + a_n(s)\,G(s,t) = 0.$$

2. Es ist:

$$G(s,s) = \frac{\partial G(s,t)}{\partial s}\bigg|_{t=s} = \cdots = \frac{\partial^{n-2} G(s,t)}{\partial s^{n-2}}\bigg|_{t=s} = 0$$

und

$$\frac{\partial^{n-1} G(s,t)}{\partial s^{n-1}}\bigg|_{t=s} = 1.$$

Auf die Gl. 1. wende man nun die Transformation $\mathfrak{T}$ an. Unter Beachtung des Hilfssatzes 6.2 erhält man:

$$\mathfrak{T}\left\{\frac{\partial^n G}{\partial s^n}\right\} + a_1(u)\,\mathfrak{T}\left\{\frac{\partial^{n-1} G}{\partial s^{n-1}}\right\} + \cdots + a_n(u)\,\mathfrak{T}\{G\} = 0.$$

Wegen der Eigenschaft 2. ergibt sich nach dem letzten Transformationsgesetz:

$$\mathfrak{T}\left\{\frac{\partial^\nu G}{\partial s^\nu}\right\} = v^\nu\,\mathfrak{T}\{G\} \quad \text{für} \quad \nu = 0, 1, \ldots, (n-1).$$

und

$$\mathfrak{T}\left\{\frac{\partial^n G}{\partial s^n}\right\} = v^n\,\mathfrak{T}\{G\} - 1\,.$$

Durch Einsetzen bekommt man:

$$\left(v^n + a_1(u)\,v^{n-1} + \cdots + a_n(u)\right)\mathfrak{T}\{G\} = 1$$

Es ist also

$$\mathfrak{T}\{G\} = h^{-1},$$

d.h. h^{-1} besitzt $G(s, t)$ als Urfunktion.

§ 7. Die lineare Differentialgleichung

Die Transformation $\mathfrak{T}$ soll nun zur Lösung der linearen Differentialgleichung

$$Y^{(n)}(s) + a_1(s)\,Y^{(n-1)}(s) + \cdots + a_n(s)\,Y(s) = F(s) \tag{7.1}$$

angewandt werden. Die $a_i(s)$ sowie $F(s)$ seien für $|s| < \delta$ analytisch.

Sei $Y(s)$ eine für $|s| < 2\delta$ analytische Lösung von 7.1. Dann ist $Y(s-t)$ für $|s| < \delta$, $|t| < \delta$ analytisch. $Y(s-t)$ wird an Stelle von $Y(s)$ in 7.1 eingesetzt. Dann kommt rechts statt $F(s)$ eine für $|s| < \delta$, $|t| < \delta$ analytische Funktion $F(s, t)$ mit $F(s, 0) = F(s)$:

$$Y^{(n)}(s-t) + a_1(s)\,Y^{(n-1)}(s-t) + \cdots + a_n(s)\,Y(s-t) = F(s,t)\,.$$

Auf diese Gleichung wird nun $\mathfrak{T}$ angewandt. Nach Hilfssatz 6.2 ist

$$\mathfrak{T}\{a_i(s)\,Y^{(i)}(s-t)\} = a_i(u)\,\mathfrak{T}\{Y^{(i)}(s-t)\}\,.$$

Es ergibt sich also:

$$\mathfrak{T}\{Y^{(n)}(s-t)\} + a_1(u)\,\mathfrak{T}\{Y^{(n-1)}(s-t)\} + \cdots + a_n(u)\,\mathfrak{T}\{Y(s-t)\} = \mathfrak{T}\{F(s,t)\} \tag{7.2}$$

Hierin drücken wir nach Hilfssatz 6.1 a) $\mathfrak{T}\{Y^{(i)}(s-t)\}$ durch $\mathfrak{T}\{Y(s-t)\}$ aus. Dies wird besonders einfach im Falle

a) $$Y(0) = Y'(0) = \cdots = Y^{(n-1)}(0) = 0\,.$$

Dann gilt:

$$\mathfrak{T}\{Y^{(i)}(s-t)\} = \mathfrak{T}\{Y(s-t)\}\cdot v^i, \qquad i = 0, 1, \ldots, n\,.$$

Da in $\mathfrak{T}\{Y(s-t)\}$ u nicht vorkommt, können wir rechts den Faktor v^i voranstellen. Mit

$$y(v) = \mathfrak{T}\{Y(s-t)\} \quad \text{und} \quad f(u, v) = \mathfrak{T}\{F(s,t)\}$$

ergibt sich die transformierte Gleichung

$$\left(v^n + a_1(u)\,v^{n-1} + \cdots + a_n(u)\right) y(v) = f(u, v)\,,$$

abgekürzt:

$$h(u, v)\,y(v) = f(u, v)\,.$$

Diese Gleichung kann innerhalb SK aufgelöst werden:

$$y = h^{-1} f\,.$$

Zu $y \in \mathsf{SK}$ kann man nun auch die zugehörige Urfunktion $Y(s-t)$ angeben: Nach Satz 5.4 gibt es zu h^{-1} eine Urfunktion $G(s,t)$, die in einer gewissen Umgebung $|s|<\delta$, $|t|<\delta'$ analytisch ist. Außerdem hat f die Urfunktion $F(s,t)$. Zu $h^{-1}f$ gehört daher nach Satz 5.3 die Volterrasche Komposition von G und F. Somit:

$$Y(s-t) = \int_t^s G(s,\sigma)\, F(\sigma,t)\, d\sigma .$$

Hierin ist zwar noch die unbekannte Funktion $F(s,t)$ enthalten, doch ergibt sich für $t=0$ wegen $F(s,0)=F(s)$:

$$Y(s) = \int_0^s G(s,\sigma)\, F(\sigma)\, d\sigma .$$

Im ganzen haben wir also den

Satz. *Berechnet man in* SK *das Inverse*

$$\left(v^n + a_1(u)\, v^{n-1} + \cdots + a_n(u)\right)^{-1} = v^{-n} + \psi_1(u)\, v^{-n-1} + \cdots$$

und geht davon durch die Ersetzungen

$$u \to s, \qquad v^{-i} \to \frac{(s-t)^{i-1}}{(i-1)!}$$

zur Urfunktion $G(s,t)$ *über, so hat man in*

$$Y(s) = \int_0^s G(s,\sigma)\, F(\sigma)\, d\sigma$$

diejenige Lösung der Differentialgleichung (7.1), *die der Anfangsbedingung*

$$Y(0) = Y'(0) = \cdots = Y^{(n-1)}(0) = 0$$

genügt. Bei Beschränkung auf reelle s *gilt dies offenbar auch für stetige* $f(s)$.

b) Es seien jetzt irgendwelche Anfangswerte $Y(0), Y'(0), \ldots, Y^{(n-1)}(0)$ vorgegeben. Nach Satz 6.1 a gilt:

$$\mathfrak{T}\{Y^{(i)}(s-t)\} = \mathfrak{T}\{Y(s-t)\}\, v^i - Y(0)\, v^{i-1} - Y'(0)\, v^{i-2} - \cdots - Y^{i-1}(0).$$

Setzt man dies in 7.2 ein und bezeichnet mit y, f wieder die Bilder von $Y(s-t)$ bzw. $F(s,t)$, so folgt:

$$\left(v^n + a_1(u)\, v^{n-1} + \cdots + a_n(u)\right) \cdot y = f + Y(0)\, v^{n-1} + Y'(0)\, v^{n-2} + \cdots + Y^{n-1}(0) +$$

$$+ a_1(u)\left\{Y(0)\, v^{n-2} + \cdots + Y^{(n-2)}(0)\right\} + \cdots + a_{n-1}(u)\, Y(0)$$

abgekürzt:

$$h\, y = f + a .$$

Das Bild von $Y(s-t)$ ist also:

$$y = h^{-1} f + h^{-1} a .$$

Es handelt sich also nur noch darum, die Urfunktion der rechten Seite zu bestimmen. Für $h^{-1}f$ steht sie bereits unter a). $h^{-1}a$ ist eine Linearkombination von Gliedern der Form

$$h^{-1}\, a(u)\, v^r .$$

Zu jedem solchen Glied läßt sich die Urfunktion angeben: Zunächst gehört nach Hilfssatz 6.3 zu $h^{-1}a(u)$ die Urfunktion $G(s,t)\,a(t)$. Für $h^{-1}a(u)v^\nu$ selbst findet man dann die Urfunktion nach Hilfssatz 6.1 b:

$$\mathfrak{T}^{-1}\{h^{-1}a(u)v^\nu\} = (-1)^\nu \frac{\partial^\nu}{\partial t^\nu}\{G(s,t)\,a(t)\}.$$

Die Anwendung dieses Satzes ist gerechtfertigt, weil h^{-1} den Grad n besitzt und $\nu \leqq (n-1)$ ist.

Setzt man die einzelnen Urfunktionen in

$$Y(s-t) = \mathfrak{T}^{-1}\{h^{-1}f\} + \mathfrak{T}^{-1}\{h^{-1}a\}$$

ein und nimmt noch die Spezialisierung $t=0$ vor, so ergibt sich der

SATZ. *Bestimmt man im Schiefkörper* SK *das Inverse* h^{-1} *und dazu die Urfunktion* $G(s,t)$, *so kann man die allgemeine Lösung von 7.1 in folgender Form erhalten:*

$$Y(s) = \int_0^s G(s,\sigma)F(\sigma)\,d\sigma + (-1)^{n-1}\frac{\partial^{n-1}G(s,t)}{\partial t^{n-1}}\Big|_{t=0} Y(0) + \cdots + G(s,0)\,Y^{(n-1)}(0) +$$
$$+ (-1)^{n-2}\frac{\partial^{n-2}}{\partial t^{n-2}}\{G(s,t)\,a_1(t)\}_{t=0}\,Y(0) + \cdots + G(s,0)\,a_1(0)\,Y^{(n-2)}(0)$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$+ G(s,0)\,a_{n-1}(0)\,Y(0).$$

Das Resultat ist bei Beschränkung auf reelle s auch für stetiges $f(s)$ gültig.

COROLLAR. *Für* $F=0$ *wird* $Y(s)$ *eine Linearkombination von*

$$G(s,0),\quad \frac{\partial G(s,t)}{\partial t}\Big|_{t=0},\ \ldots,\quad \frac{\partial^{n-1}G(s,t)}{\partial t^{n-1}}\Big|_{t=0}.$$

Diese Funktionen bilden also ein Fundamentalsystem der homogenen Gleichung.

Das Inverse

$$h^{-1} = v^{-n} + \psi_1(u)\,v^{-n-1} + \psi_2(u)\,v^{-n-2} + \cdots$$

liefert demnach ein Fundamentalsystem in Form der folgenden Reihen:

$$\frac{s^{n-1}}{(n-1)!} + \psi_1(s)\frac{s^n}{n!} + \psi_2(s)\frac{s^{n+1}}{(n+1)!} + \cdots$$
$$\frac{s^{n-2}}{(n-2)!} + \psi_1(s)\frac{s^{n-1}}{(n-1)!} + \psi_2(s)\frac{s^n}{n!} + \cdots$$
$$1 + \psi_1(s)\frac{s}{1!} + \psi_2(s)\frac{s^2}{2!} + \cdots.$$

Ähnlich gestaltet sich die Auflösung der adjungierten Differentialgleichung

$$(-1)^n Y^{(n)}(t) + (-1)^{n-1}\left(a_1(t)\,Y(t)\right)^{(n-1)} + \cdots + a_n(t)\,Y(t) = F(t)$$

mittels $\mathfrak{T}$. Man nehme zuerst die Modifikation

$$(-1)^n Y^{(n)}(t-s) + (-1)^{n-1}\left(a_1(t)\,Y(t-s)\right)^{(n-1)} + \cdots + a_n(t)\,Y(t-s) = F^*(s,t)$$

vor. Der Kürze halber beschränken wir uns auf diejenige Lösung $Y(t)$ mit $Y(0) = \cdots = Y^{(n-1)}(0) = 0$. Nach den Hilfssätzen 6.1 und 6.3 ergibt dann die

Anwendung von $\mathfrak{T}$:

$$\mathfrak{T}\{Y(t-s)\}\,v^n + \mathfrak{T}\{Y(t-s)\}\,a_1(u)\,v^{n-1} + \cdots + \mathfrak{T}\{Y(t-s)\}\,a_n(u) = \mathfrak{T}\{F^*\}$$

also:

$$\mathfrak{T}\{Y(t-s)\} = \mathfrak{T}\{F^*\}\,h^{-1}.$$

Es folgt:

$$Y(t-s) = \int_t^s F^*(s,\sigma)\,G(\sigma,t)\,d\sigma$$

und für $s=0$:

$$Y(t) = -\int_0^t G(\sigma,t)\,F(\sigma)\,d\sigma.$$

h^{-1} tritt hier als Rechtsinverses auf, während es oben als Linksinverses fungierte. Dies bringt den Übergang von $G(s,t)$ zu $-G(t,s)$ mit sich.

Bemerkungen. 1. Die Transformation $\mathfrak{T}$ kann auch zur Lösung der Integrodifferentialgleichung

$$Y^{(n)}(s) + a_1(s)\,Y^{(n-1)}(s) + \cdots + a_n(s)\,Y(s) + \int_0^s K(s,\sigma)\,Y(\sigma)\,d\sigma = F(s),$$

bei der $a_i(s)$, $K(s,t)$, $F(s)$ analytische Funktionen sind, angewandt werden. Wir beschränken uns dabei wieder auf diejenige Lösung, die der homogenen Anfangsbedingung genügt. Zur Lösung mittels $\mathfrak{T}$ ersetze man in $Y(s)$ das Argument durch $s-t$, ferner die untere Integrationsgrenze 0 durch t. Auf die entstehende Gleichung

$$Y^{(n)}(s-t) + a_1(s)\,Y^{(n-1)}(s-t) + \cdots + a_n(s)\,Y(s-t) + \\ + \int_t^s K(s,\sigma)\,Y(\sigma-t)\,d\sigma = F^*(s,t),$$

die für $t=0$ in die gegebene Gleichung übergeht, wende man $\mathfrak{T}$ an, was eine einfache algebraische Gleichung

$$h\,\mathfrak{T}\{Y(s-t)\} = \mathfrak{T}\{F^*\}$$

in bezug auf SK ergibt, bei der

$$h = v^n + a_1(u)\,v^{n-1} + \cdots + a_n(u) + \sum_{\nu=0}^{\infty} \varrho_\nu(u)\,v^{-\nu-1}$$

ist, wenn $K(s,t)$ die Entwicklung

$$K(s,t) = \sum_{\nu=0}^{\infty} \varrho_\nu(s)\,\frac{(s-t)^\nu}{\nu!}$$

besitzt.

Die Lösung in SK ist:

$$\mathfrak{T}\{Y(s-t)\} = h^{-1}\,\mathfrak{T}\{F^*\}.$$

Man kann zeigen, daß h^{-1} Funktionselement ist. Gehört dazu die Urfunktion $G(s,t)$, so ist nach dem „Faltungssatz"

$$Y(s-t) = \int_t^s G(s,\sigma)\,F^*(\sigma,t)\,d\sigma,$$

also

$$Y(s) = \int_0^s G(s,\sigma)\, F(\sigma)\, d\sigma .$$

Dieses Resultat ist bei Beschränkung auf reelle s auch für stetiges $F(s)$ gültig.

2. Es möge noch angedeutet werden, wie sich die hier angegebene Methode mit dem Kalkül von MIKUSINSKI in Zusammenhang bringen läßt.

Die Grundlage für diesen Kalkül ist bekanntlich ein gewisser kommutativer Körper **MK**, dessen Produkt mittels der Faltung zweier Funktionen eingeführt wird. Dieser bietet den Vorteil, daß er sowohl Funktionen wie Operatoren umfaßt. Beispielsweise läßt sich die Differentialgleichung

$$y^{(n)} + a_1\, y^{(n-1)} + \cdots + a_n\, y = f(t),$$

deren Koeffizienten *konstant* sind, direkt als Gleichung in **MK** schreiben:

$$s^n\{y\} + \alpha_1\, s^{n-1}\{y\} + \cdots + \alpha_n\{y\} = \{f\}.$$

Wie bei MIKUSINSKI bedeutet hierbei s den Differentiationsoperator, der als das Inverse des Integrationsoperators definiert ist. Ferner ist:

$$\alpha_i = \frac{\{a_i\}}{\{1\}}.$$

Leider ist jedoch der Mikusinskische Kalkül auf Differentialgleichungen mit *konstanten* Koeffizienten beschränkt. Nach Vorhergehendem läßt sich aber leicht ein „Operatorkörper" angeben, der auch den Fall analytischer Koeffizienten erfaßt: Man gehe nämlich aus von der Menge der komplexen Funktionen $K(s,t)$, die um 0 analytisch sind, definiere für sie das Produkt als Volterrasche Komposition

$$\{K_1\}\{K_2\} = \int_t^s K_1(s,\sigma)\, K_2(\sigma,t)\, d\sigma$$

und hat dann, wenn man noch die gewöhnliche Summe als erste Verknüpfung nimmt, einen gewissen Ring $\mathfrak{V}$. Nach §4 läßt sich $\mathfrak{V}$ isomorph in den Schiefkörper **SK** abbilden. Daher ist $\mathfrak{V}$ in einen Körper einbettbar.

Man kann nun ähnlich wie MIKUSINSKI vorgehen:

Die Funktion $\{1\}$ wird als „Integrationsoperator l" eingeführt. Nach Definition des Produkts für $\mathfrak{V}$ ist dann:

$$l^n = \left\{\frac{(s-t)^{n-1}}{(n-1)!}\right\}.$$

Der „Operator der Differentiation" kann nun definiert werden durch

$$D = l^{-1}.$$

Für eine Funktion $y(s-t)$ von $s-t$ allein gilt dann offenbar:

$$D\{y(s-t)\} = \{y'(s-t)\} + D\{y(0)\}.$$

Setzt man in Analogie zu MIKUSINSKI

$$\{a(s)\}\, D = \alpha,$$

so erhält man für die der homogenen Anfangsbedingung genügende Lösung von

$$y^{(n)}(s-t) + a_1(s)\, y^{(n-1)}(s-t) + \cdots + a_n(s)\, y(s-t) = f^*(s,t)$$

die „Operatorgleichung"

$$(D^n + \alpha_1 D^{n-1} + \cdots + \alpha_n)\, \{y(s-t)\} = \{f^*\},$$

die im „Operatorkörper" die Lösung

$$\{Y(s-t)\} = (D^n + \alpha_1 D^{n-1} + \cdots + \alpha_n)^{-1}\, \{f^*\}$$

besitzt. Unter Zuhilfenahme von Satz 5.4 läßt sich nun beweisen, daß das Inverse

$$(D^n + \alpha_1 D^{n-1} + \cdots + \alpha_n)^{-1}$$

mit der bereits benutzten Funktion $G(s,t)$ identisch ist, so daß also die Lösung im Operatorkörper auch in der Form

$$\{y(s-t)\} = \{G(s,t)\}\, \{f^*(s,t)\}$$

geschrieben werden kann. Wegen der getroffenen Produktdefinition ist dies gleichbedeutend mit:

$$y(s-t) = \int_t^s G(s,\sigma)\, f^*(\sigma,t)\, d\sigma,$$

was für $t=0$ die Lösung

$$Y(s) = \int_0^s G(s,\sigma)\, f(\sigma)\, d\sigma$$

ergibt.

3. Der Schiefkörper SK enthält einen Ring, der zum Ring $\mathfrak{R}_D$ der linearen Differentialoperatoren

$$a_0(s)\, D^n + a_1(s)\, D^{n-1} + \cdots + a_n(s)$$

mit analytischen $a_i(s)$ isomorph ist. Jedoch stellt SK nicht die minimale Erweiterung von $\mathfrak{R}_D$ zum Körper dar. Letztere läßt sich unter Zuhilfenahme des Einbettungssatzes von O. ORE als Quotientenkörper über $\mathfrak{R}_D$ konstruieren. Der Zusammenhang zwischen diesen beiden Möglichkeiten der Einbettung wird in einer folgenden Arbeit erörtert.

Literatur

BERG, E. J.: Rechnung mit Operatoren nach OLIVER HEAVISIDE. München u. Berlin 1932.

BROMWICH, T. J. I. A.: Normal coordinates in dynamical systems. Proc. Lond. Math. Soc. **15**, 401 (1916).

CARSON, J. R.: Electric circuit theory and the operational calculus. Bell Technical J. **4**, 685 (1925); **5**, 50 (1926).

CHURCHILL, R. V.: Operational Mathemasics. New York-Toronto-London 1958.

DOETSCH, G.: Die Integrodifferentialgleichungen vom Faltungstypus. Math. Annalen **89**, 192 (1923).

— Die Anwendung der Funktionaltransformation in der Theorie der Differentialgleichungen und die symbolische Methode (Operatorenkalkül). Jahresber. dtsch. Math. Vereinigung **43**, 238 (1934).

— Handbuch der Laplace-Transformation. Basel u. Stuttgart 1950–1956.

DROBOT, S., et J. G. MIKUSINSKI: Sur l'unicité des solutions de quelques équations differentielles dans les espaces abstraits. Studia Mathematica **11**, 38–40 (1950).

HEAVISIDE, O.: Electromagnetic Theory. London 1899.

HUMBERT, P.: Formules nouvelles pour le calcul symbolique. Bull. Soc. Math. France **65**, 119 (1937).

KAMKE, E.: Differentialgleichungen I, 3. Aufl. Leipzig 1944.

KOPPENFELS, W. v.: Der Faltungssatz und seine Anwendung bei der Integration linearer Differentialgleichungen mit konstanten Koeffizienten. Math. Ann. **105**, 694 (1931).

MIKUSINSKI, J. G.: Operatorenrechnung. Berlin 1957.

— Sur les fondements du calcul operatoire. Studia Mathematica **11**, 41–70 (1950).

PICONE, M.: Nouvelles methodes du recherche pour la détermination des intégrales des équations linéaires aux dérivées partielles. Ann. du la Soc. Pol. de Math. **19**, 36–61 (1946).

VAHLEN, K. TH.: Über den Heaviside-Kalkül. Z. angew. Math. u. Mechanik **13**, 283 (1933).

WAGNER, K. W.: Operatorenrechnung und Laplacesche Transformation. Leipzig 1950.

Gießen, Mathematisches Institut der Universität

(Eingegangen am 3. August 1960)

KÜNNETH, H.
Math. Zeitschr. 76, 199–208 (1961)

Zu einem Satz von Hjelmslev über monotone Bogen

Von
HERMANN KÜNNETH

In seiner grundlegenden Arbeit über monotone Folgen und Bogen bringt HJELMSLEV[1]) im *Theorem* 28 auch den wichtigen Satz: „Der Richtungskegel einer monotonen Kurve ist selbst monoton oder mindestens aus einer endlichen Zahl monotoner Teilbogen zusammengesetzt." Der Richtungskegel wird von den durch einen beliebigen Punkt 0 gehenden Parallelen zu den vorderen Halbtangenten gebildet. Das Theorem wird aus Sätzen über monotone Punktfolgen abgeleitet. Eine Punktfolge p_i $(i=1, 2, \ldots)$ im euklidischen R^n heißt *monoton*, wenn *jedes* n-dimensionale Simplex $p_{i_0} p_{i_1} \ldots p_{i_n}$ $(i_0<i_1<\cdots<i_n)$ die gleiche Orientierung hat. Der Richtungskegel der Folge wird entsprechend wie oben erklärt, wobei an Stelle der Halbtangenten die von p_i ausgehenden Halbgeraden $p_i p_{i+1}$ treten. Nach Theorem 16 ist dieser Richtungskegel schließlich monoton, worauf sich der Beweis von Theorem 28 stützt. Theorem 16 ist aber nur für den R^3 bewiesen, und auf S. 58 heißt es nur „le théorèm 16 peut être facilement généralisé a l'espace R^n". Dazu müßte das Theorem 5 auf den R^n verallgemeinert werden, das für den R^2 aussagt, daß eine unendliche beschränkte Punktfolge p_i, die nicht schließlich in einer Geraden liegt, schließlich monoton ist, wenn 1. drei aufeinanderfolgende Punkte p_i, p_{i+1}, p_{i+2}, falls sie nicht in einer Geraden liegen, immer die gleiche Orientierung haben, und 2. die Halbgeraden $p_{i+1} p_i$ eine Halbgerade zum Limes haben.

Zum Beweis des Theorems 28 ist daher nötig die Verallgemeinerung der Theoreme 5 und 16 auf den R^n, und hiezu benötigt man eine Verallgemeinerung auf den R^n von Theorem 4, nach dem die Halbgeraden $p_s p_r$ $(s>r)$ einer beschränkten monotonen Punktfolge im R^2 eine Halbgerade zum Limes haben. Dies ist in Theorem 22 nur geschehen für den Fall der Hyperebenen $(p_{i_1}, p_{i_2} \ldots p_{i_n})$[2]) $(i_1<i_2<\cdots<i_n)$. Im übrigen heißt es nur (S. 57): „On peut définir d'une manière analogue le demi-espaoe osculateur et on peut aller plus loin en considérant l'espace osculateur R^p à un nombre quelconque p de dimensions, $p<n$."

Hier sollen nun die Lücken im Beweis des wichtigen Theorems 28 ausgefüllt und dazu die Verallgemeinerungen der Theoreme 5, 16 und 22 gebracht werden.

[1]) „Introduction à la théorie des suites monotones." Overs. ov. d. kong. Danske Vidensk. selskabs Forhandlinger. S. 3–74. (1914).

[2]) Mit $(Q_1, Q_2, \ldots, Q_\nu)$ werde der kleinste lineare Unterraum von R^n bezeichnet, der die linearen Unterräume $Q_1, Q_2, \ldots, Q_\nu$ enthält, mit $[Q_1; q)$ der von Q_1 begrenzte abgeschlossene Halbraum, der den Punkt q enthält, mit $p_{i_0} p_{i_1} \ldots p_{i_r}$ das r-dimensionale orientierte Simplex (r-Simplex), das durch die Eckpunkte $p_{i_0}, p_{i_1}, \ldots, p_{i_r}$ in dieser Reihenfolge bestimmt wird. Ein r-dimensionaler linearer Unterraum von R^n $(0 \leqq r \leqq n-1)$ werde als r-Ebene bezeichnet, eine $(n-1)$-Ebene auch als Hyperebene.

1. $p_1, p_2, p_3 \ldots$ sei eine unendliche Punktfolge F im euklidischen R^n. Die von $r+1$ linear unabhängigen Punkte $p_{i_1}, p_{i_2} \ldots p_{i_{r+1}}$ ausgespannten r-Ebenen heißen *r-Sekanten* der Folge F $(0 \leqq r < n)$, und *r-Schmieg-Sekanten*, wenn die $r+1$ Punkte p_{i_ϱ} aufeinanderfolgende Punkte der Folge sind, also $i_{\varrho+1} = i_\varrho + 1$. Die r-Schmieg-Sekanten werden bezeichnet mit $S_i^r = (p_i, p_{i+1} \ldots p_{j+r})$, $i \geqq 1$, $0 \leqq r \leqq n$; $S_i^0 = p_i$, $S_i^n = R^n$.

Das r-Simplex $p_i p_{i+1} \ldots p_{i+r}$ werde mit $\mathfrak{S}_i^r$ bezeichnet.

2. Allen r-Sekanten der Folge $(1 \leqq r \leqq n-1)$ sei eine Orientierung beigelegt. Dadurch zerfallen die in ihnen liegenden r-Simplexe in zwei Äquivalenzklassen, die Klasse der positiven und der negativen r-Simplexe. *Die Klasseneinteilung sei so getroffen, daß das r-Simplex $p_{i_1} p_{i_2} \ldots p_{i_{r+1}}$ in $(p_{i_1}, p_{i_2}, \ldots, p_{i_{r+1}})$ der positiven Klasse angehört.*

3. HJELSMLEV hatte bewiesen (a.a.O., Theorem 21), daß die Zentralprojektion einer beschränkten monotonen Punktfolge aus einem beliebigen Punkt q des R^n in eine Hyperebene H^{n-1} eine schließlich monotone Punktfolge in H^{n-1} ist.

Eine Verallgemeinerung davon liefert den

SATZ 1. *Die Projektion einer beschränkten monotonen Punktfolge aus einer k-Ebene R^k $(0 \leqq k \leqq n-2)$ in eine zu R^k fremde $(n-k-1)$-Ebene H^{n-k-1} ist (falls beschränkt) eine schließlich monotone Punktfolge (in H^{n-k-1}).*

BEWEIS. Der Satz ist für $k=0$ von HJELMSLEV bewiesen. Wir nehmen an, er sei richtig für $k < k_0 \leqq n-2$. Die Punktfolge F werde aus der k_0-Ebene R^{k_0} in die dazu fremde $(n-k_0-1)$-Ebene H^{n-k_0-1} projiziert. Die Projektion von F sei F'.

Es sei R^{k_0-1} irgend eine in R^{k_0} enthaltene (k_0-1)-Ebene und q ein Punkt in R^{k_0}, $q \notin R^{k_0-1}$, also $R^{k_0} = (R^{k_0-1}, q)$; ferner sei $H^{n-k_0} = (H^{n-k_0-1}, q)$.

F' kann man dann auch erhalten, indem man zunächst F aus R^{k_0-1} in H^{n-k_0} projiziert, die Projektion ist nach Induktionsannahme eine schließlich monotone Punktfolge F'', diese projiziert man nun aus q in H^{n-k_0-1}, was wieder eine schließlich monotone Punktfolge ergibt.

4. Die Verallgemeinerung von Theorem 22 erhält man nun so: Jede beschränkte monotone Punktfolge hat einen eindeutig bestimmten Limespunkt p (HJELMSLEV, a.a.O., S. 41). Gäbe es nun zwei Folgen gegen p konvergierender $(k+1)$-tupel $p_{i_0}, p_{i_1}, \ldots, p_{i_k}$ und $p_{i'_0}, p_{i'_1}, \ldots, p_{i'_k}$, so daß die orientierten k-Ebenen $(p_{i_0}, p_{i_1}, \ldots, p_{i_k})$ bzw. $(p_{i'_0}, p_{i'_1}, \ldots, p_{i'_k})$ eine orientierte k-Ebene α bzw. α' zum Limes haben $(1 \leqq k \leqq n-1)$, wobei α und α' verschiedene k-Ebenen sind oder dieselbe k-Ebene, aber mit verschiedener Orientierung, so ließe sich immer eine zu α und α' fremde $(n-k-1)$-Ebene R^{n-k-1} finden und $n-k$ unabhängige Punkte $q_1, q_2 \ldots q_{n-k}$ in R^{n-k-1}, so daß die n-Simplexe $q_1 q_2 \ldots q_{n-k} p_{i_0} p_{i_1} \ldots p_{i_k}$ und $q_1 q_2 \ldots q_{n-k} p_{i'_0} p_{i'_1} \ldots p_{i'_k}$ verschieden orientiert sind und daher die Projektion von F aus R^{n-k-1} in eine dazu fremde k-Ebene nicht schließlich monoton ist, in Widerspruch zu Satz 1.

Alle k-Sekanten einer beschränkten monotonen Punktfolge haben daher eine eindeutig bestimmte orientierte k-Ebene zum Limes.

5. Die Verallgemeinerung von Theorem 5 lautet nun so:

SATZ 2. *Eine unendliche Punktfolge $F=p_1, p_2 \ldots$ im R^n ist schließlich monoton, wenn sie folgende Forderungen erfüllt:*

FORDERUNG 1. *$n+1$ aufeinanderfolgende Punkte sind linear unabhängig und die von ihnen ausgespannten n-Simplexe haben alle dieselbe Orientierung.*

FORDERUNG 2. *Alle orientierten k-Schmieg-Sekanten $(0 \leqq k \leqq n-1)$ haben eine orientierte k-Ebene zum Limes.*

Mit Forderung 2 ist gleichbedeutend

FORDERUNG 2'. *Alle Halb-k-Ebenen $[S_i^{k-1}; p_{i+k})$ haben eine Halb-k-Ebene zum Limes $(1 \leqq k \leqq n-1)$.*

Aus Forderung 1 und 2 folgt, daß auch die Halbräume $[S_i^{n-1}; p_{i+n})$ einen Halbraum von R^n zum Limes haben[3]).

Von den beiden Halb-k-Ebenen, in die S_i^k durch S_i^{k-1} $(1 \leqq k \leqq n)$ geteilt wird, werde $[S_i^{k-1}; p_{i+k})$ als die *„positive"* Halb-k-Ebene bezeichnet und alle in $[S_i^{k-1}; p_{i+k})$ liegenden Punkte als *„oberhalb"* S_i^{k-1} in S_i^k gelegen.

Nach Forderung 2 existiert $\lim_{i\to\infty} S_i^k = E^k$, wobei E^k eine orientierte k-Ebene ist $(0 \leqq k \leqq n-1)$, $\lim_{i\to\infty} p_i = E^0 = p$.

Es gibt also ein i_0 und ein $\varepsilon > 0$, so daß alle p_i mit $i > i_0$ in einer konvexen ε-Umgebung $V(p)$ von p liegen. Die Punktfolge F ist daher beschränkt.

Es sei $E_+^k = \lim_{i\to\infty} [S_i^{k-1}; p_{i+k})$ $(1 \leqq k \leqq n)$.

Als eine zu F zugehörige *„Inzidenzkette"* (IK) von k-Ebenen $(k=0, 1, \ldots, n-1)$ werde jede Folge $U^0 \subset U^1 \subset U^2 \ldots \subset U^{n-1}$ bezeichnet, für die folgendes gilt:

Für alle $i > i_0$ ist 1. $U^k \cap S_i^{n-k-1} = \emptyset$ $(0 \leqq k \leqq n-1)$, 2. $U^0 \in [S_i^{n-1}; p_{i+n})$; $U^1 \cap S_i^{n-1} = x_i^{n-1} \in [S_i^{n-2}; p_{i+n-1})$; $U^1 \cap E^{n-1} = e^{n-1} \in E_+^{n-1}$; $U^k \cap S_i^{n-k} = x_i^{n-k} \in [S_i^{n-k-1}; p_{i+n-k})$; $U^k \cap E^{n-k} = e^{n-k} \in E_+^{n-k}$ $(1 \leqq k \leqq n-1)$, 3. Es gibt eine zu U^{k-1} fremde konvexe Umgebung $V(e^{n-k})$ von e^{n-k} in U^k, so daß $x_i^{n-k} \in V(e^{n-k})$.

Die an die Inzidenzkette U^k $(k=0, 1, \ldots, n-1)$ gestellten Forderungen können gemäß der Forderung 2' des Satzes 2 erfüllt werden.

Für die weiteren Betrachtungen sei immer $i > i_0$ angenommen.

Projiziert man die k-Simplexe $(1 \leqq k \leqq n-1)$ einer Äquivalenzklasse von S_i^k aus U^{n-k-1} in S_j^k oder E^k, so gehören die Projektionen der k-Simplexe auch alle derselben Äquivalenzklasse in S_j^k bzw. E^k an. Das folgt aus der durch die Projektion erzeugten eineindeutigen stetigen Abbildung und daraus, daß Simplexe derselben Äquivalenzklasse und nur solche stetig ineinander übergeführt werden können. So läßt sich durch die Projektion aus U^{n-k-1} die Orientierung von S_i^k auf S_j^k bzw. E^k übertragen.

---

[3]) HJELMSLEV verlangt in Theorem 5 bis, bzw. 17 und 23 nur, daß die Punktfolge nicht schließlich in einer Hyperebene liegt und daß $[S_i^{n-1}; p_{i+n})$ für $i \to \infty$ einen Halbraum zum Limes hat. Das genügt aber nicht, wie das einfache Gegenbeispiel der Punktfolge $p_i\left(x = \frac{(-1)^i}{i};\ y = \frac{1}{i}\right)$ für den R^2 beweist. In R^3 genügt die Punktfolge $p_i\left(x = \frac{1}{i}\cos\frac{2\pi i}{3};\ y = \frac{1}{i}\sin\frac{2\pi i}{3};\ z = \frac{1}{i}\right)$ außerdem sogar noch der Forderung 1 ohne schließlich monoton zu sein.

Mit Forderung 2 ist daher auch gleichbedeutend

FORDERUNG 2''. *Die Projektionen der k-Simplexe $\mathfrak{S}_i^k$ $(i=i_0, i_0+1, \ldots)$ aus U^{n-k-1} in S_j^k oder E^k gehören alle derselben (nämlich der positiven) Äquivalenzklasse in S_j^k bzw. E^k an* $(1 \leqq k \leqq n-1)$.

6. Zum Beweis von Satz 2 benötigen wir einiger Hilfssätze.

HILFSSATZ 1. *Projiziert man die Punkte einer Punktfolge F, die den Forderungen 1 und 2 des Satzes 2 genügt, aus U^{n-r-1} in S_i^r, so erhält man eine Punktfolge F'; die ebenfalls den Forderungen 1 und 2 genügt, wobei n durch r zu ersetzen ist und U^ϱ durch $S_i^r \cap U^\varrho = U'^{r+\varrho-n}$* $(n-r \leqq \varrho \leqq n-1)$.

BEWEIS. 1. Da für F Forderung 2'' erfüllt ist, sind alle Projektionen der r-Simplexe $\mathfrak{S}_j^r$ $(j > i_0)$ aus U^{n-r-1} in S_i^r positive r-Simplexe in S_i^r; daraus folgt, daß Forderung 1 für die Folge F' erfüllt ist.

2. Da $S_j^\varrho \cap U^{n-r-1} = \emptyset$ für $\varrho \leqq r$, ist $(U^{n-r-1}, S_j^\varrho) \cap S_i^r = T_j^\varrho$ eine ϱ-Schmieg-Sekante von F' und, da $\lim\limits_{j\to\infty} T_j^\varrho = \lim\limits_{j\to\infty} (U^{n-r-1}, S_j^\varrho) \cap S_i^r = (U^{n-r-1}, E^\varrho) \cap S_i^r$, haben alle ϱ-Schmieg-Sekanten von F' $(0 \leqq \varrho < r)$ einen Limes. Es ist noch zu zeigen, daß sie eine orientierte ϱ-Ebene zum Limes haben.

3. $x' = (U^{n-r-1}, x) \cap S_i^r$ sei die Projektion eines Punktes x aus U^{n-r-1} in S_i^r.

$x'' = (U'^{r-\sigma-1}, x') \cap T_j^\sigma$ sei die Projektion des Punktes x' aus $U'^{r-\sigma-1}$ in T_j^σ $(1 \leqq \sigma \leqq r-1)$.

Es ist $\dim(S_i^r \cap U^{n-\sigma-1}) = r-\sigma-1$ und $\dim\big((S_i^r \cap U^{n-\sigma-1}), U^{n-r-1}\big) = n-\sigma-1$, da $S_i^r \cap U^{n-r-1} = \emptyset$; also, da $U^{n-r-1} < U^{n-\sigma-1}$, $x'' = \big((S_i^r \cap U^{n-\sigma-1}), U^{n-r-1}, x\big) \cap T_j^\sigma = (U^{n-\sigma-1}, x) \cap T_j^\sigma$, d.h. x'' ist die Projektion von x aus $U^{n-\sigma-1}$ in T_j^σ. Man erhält daher die gleichen σ-Simplexe, ob man die σ-Simplexe $\mathfrak{S}_j^\sigma$ $(0 \leqq \sigma \leqq r-1)$ aus $U^{n-\sigma-1}$ in T_j^σ projiziert oder erst aus U^{n-r-1} in S_i^r und dann aus $U'^{r-\sigma-1}$ in T_j^σ.

Daher gehören die Projektionen der σ-Simplexe $\mathfrak{S}_j^\sigma$ $(0 \leqq \sigma \leqq r-1)$ aus $U^{n-\sigma-1}$ in das orientierte T_j^σ alle der gleichen (positiven) Klasse an; dasselbe gilt auch, wenn man S_j^σ durch E^σ, also T_j^σ durch $(U^{n-r-1}, E^\sigma) \cap S_i^r$ ersetzt. So ist auch Forderung 2'' (und damit 2) für F' erfüllt.

7. HILFSATZ 2. *Wenn für eine Punktfolge F: $p_1, p_2 \ldots$ die Forderungen des Satzes 2 erfüllt sind, so gibt es ein i_0, so daß p_k oberhalb S_i^{n-1} liegt für $k \geqq i+n$ und $i > i_0$* $(n \geqq 1)$.

Genauer wird folgendes behauptet: a) Für $i > i_0$, $k \geqq i+n$ und beliebige IK bezüglich F liegt p_{i+n} in dem n-Simplex Q_i, das begrenzt wird von den $(n-1)$-Ebenen (U^{n-2}, S_{i+n-1}^0), (U^{n-2}, p_{i+n-1}), $(U^{n-3}, S_{i+n-2}^1) \ldots (U^0, S_{i+1}^{n-2})$, von U^{n-1} und S_i^{n-1}, oder wenn U^{-1} und S^{-1} die leere Menge bedeutet, von den $(n-1)$-Ebenen $(U^{n-r-1}, S_{i+n-r}^{r-1})$ $(0 \leqq r \leqq n)$.

b) Es ist $Q_{i'} < Q_i$ für $i' > i$.

BEWEIS. Der Hilfssatz ist richtig für $n=1$. Wir nehmen an, er sei richtig für den R^{n-1}, und beweisen ihn für den R^n.

Man projiziert F aus U^0 in S_i^{n-1}. p'_ν sei die Projektion von p_ν $(\nu = i, i+1 \ldots)$. Dann bilden gemäß Hilfssatz 1 die Punkte $p'_i = p_i$, $p'_{i+1} = p_{i+1} \ldots p'_{i+n-1} =$

$p_{i+n-1}, p'_{i+n}, \ldots, p'_k \ldots$ eine Punktfolge F', die Forderung 1 und 2 von Satz 2 erfüllt. $U^\varrho \cap S_i^{n-1}$ $(1 \leqq \varrho \leqq n-1)$ bildet eine IK für F'. Die Ecken von Q_i sind: $S^\varrho_{i+n-\varrho-1} \cap U^{n-\varrho}$ $(0 \leqq \varrho \leqq n) = x^\varrho_{i+n-\varrho-1}$. Denn es ist $S^\varrho_{i+n-\varrho-1} \cap U^{n-\varrho}$ enthalten in $(S^{r-1}_{i+n-r}, U^{n-r-1})$ für $r < \varrho$, da $U^{n-\varrho} < U^{n-r-1}$, und für $r > \varrho$, da $S^\varrho_{i+n-\varrho-1} < S^{r-1}_{i+n-r}$. Dagegen ist $S^r_{i+n-r-1} \cap U^{n-r} \not\subset (S^{r-1}_{i+n-r}, U^{n-r-1})$, da $S^r_{i+n-r-1} \cap U^{n-r-1} = \emptyset$ und $U^{n-r} \cap S^{r-1}_{i+n-r} = \emptyset$. Die Ecke $x^r_{i+n-r-1}$ liegt aber der in $(S^{r-1}_{i+n-r}, U^{n-r-1})$ liegenden Seite von Q_i gegenüber.

a) Da nach Induktionsannahme Hilfssatz 2 für F' in S_i^{n-1} anwendbar ist, liegt p'_{i+n} in demjenigen $(n-1)$-Simplex Q'_{i+1} in S_i^{n-1}, das begrenzt wird von den $(n-2)$-Ebenen $((U^{n-r-1} \cap S_i^{n-1}), S^{r-1}_{i+n-r})$ $(r = 0, 1, \ldots, n-1)$ mit den Ecken $x^\varrho_{i+n-\varrho-1}$ $(0 \leqq \varrho \leqq n-1)$. Q'_{i+1} hat also als Seiten die Durchschnitte von S_i^{n-1} mit den übrigen Seiten von Q_i und ist eines der $(n-1)$-dimensionalen Randsimplexe von Q_i, also $Q'_{i+1} = Q_i \cap S_i^{n-1}$. p_{i+n} liegt zwischen U^0 und p'_{i+n}, also in Q_i.

b) Q_{i+1} wird begrenzt von den $(n-1)$-Ebenen $(U^{n-r-1}, S^{r-1}_{i+n-r+1})$ $(0 \leqq r \leqq n)$ $\overline{Q}_{i+1}$ sei das n-Simplex, das begrenzt wird von den $(n-1)$-Ebenen $(U^{n-r-1}, S^{r-1}_{i+n-r+1})$ $(0 \leqq r \leqq n-1)$ und von S_i^{n-1}. $\overline{Q}_{i+1}$ hat also dieselben Begrenzungshyperebenen wie Q_{i+1}, nur statt S_{i+1}^{n-1} die $(n-1)$-Schmieg-Sekante S_i^{n-1}.

Q'_{i+2} in S_i^{n-1} wird begrenzt von den Durchschnitten der Hyperebenen $(U^{n-r-1}, S^{r-1}_{i+n-r+1})$ $(0 \leqq r \leqq n-1)$ mit S_i^{n-1}, da alle diese Hyperebenen durch das Projektionszentrum U^0 gehen.

Q'_{i+2} ist die in S_i^{n-1} liegende Seite von $\overline{Q}_{i+1}$ und es ist nach Induktionsannahme $Q'_{i+2} < Q'_{i+1}$, also auch $\overline{Q}_{i+1} < Q_i$.

$S_i^{n-1}, S_{i+1}^{n-1}, (S_{i+1}^{n-2}, U^0)$ gehören demselben Hyperebenenbüschel mit der Achse S_{i+1}^{n-2} an. S_i^{n-1} und (S_{i+1}^{n-2}, U^0) sind Begrenzungshyperebenen von Q_i. Da p_{i+n} in Q_i liegt, so ist $S_{i+1}^{n-1} \cap U^1 = x_{i+1}^{n-1}$ auf U^1 zwischen $S_i^{n-1} \cap U^1 = x_i^{n-1}$ und U^0. Weiter folgt daraus, daß $Q_{i+1} < Q_i$.

Aus $Q_{i+1} < Q_i$ und $p_{i+n} \in Q_i$ erhält man: $Q_i < Q_{i+1} < Q_{i+2} < \cdots$ und $p_{\nu+n}$ $(\nu \geqq i) \in Q_\nu$, woraus folgt: $p_k \in Q_i$ $(k \geqq i+n)$.

Zusatz. Außerdem hat sich aus dem Beweis ergeben, daß auf U^1 die Punkte $x_i^{n-1}, x_{i+1}^{n-1}, x_{i+2}^{n-1} \ldots U^0$ in dieser Reihenfolge aufeinanderfolgen.

8. *r-Sekanten-Hüllen* von F $(0 \leqq r \leqq n-1)$. Die 0-Sekanten-Hülle $H(p_{i_1}, p_{i_2} \ldots p_{i_r})$ der endlich vielen Punkte $p_{i_1}, p_{i_2} \ldots p_{i_r}$ von F bezüglich einer IK, auch als *Punkthülle* bezeichnet, ist gleich der konvexen Hülle dieser Punkte im R^n. Sie ist enthalten in $V(E^0)$ (5)[4] $H(p_{i_1}, p_{i_2}, \ldots, p_{i_r}) < (p_{i_1}, p_{i_2}, \ldots, p_{i_r})$.

Sind $S^1_{i_1}, S^1_{i_2} \ldots S^1_{i_r}$ irgend welche 1-Schmieg-Sekanten von F und ist $S^1_{i_\varrho} \cap U^{n-1} = x^1_{i_\varrho}$ $(\varrho = 1, 2, \ldots, r)$, ist ferner $H(x^1_{i_1}, x^1_{i_2} \ldots x^1_{i_r})$ die Punkthülle der $x^1_{i_\varrho}$ in $V(e^1)$ (5) und $H(p_{j_1}, p_{j_2} \ldots p_{j_s})$ die Punkthülle der mit den $S^1_{i_\varrho}$ $(\varrho = 1, 2, \ldots, r)$ inzidenten Punkte p_{j_σ} von F $(\sigma = 1, 2, \ldots, s)$, so sei unter der *1-Sekanten-Hülle* $H(S^1_{i_1}, S^1_{i_2} \ldots S^1_{i_r})$ der $S^1_{i_1}, S^1_{i_2} \ldots S^1_{i_r}$ bezüglich der IK verstanden: die Menge der *Geraden*, welche Punkte von $H(x^1_{i_1}, x^1_{i_2} \ldots x^1_{i_r})$ mit Punkten von $H(p_{j_1}, p_{j_2} \ldots p_{j_s})$ verbinden.

[4]) Zahlen in Klammern verweisen auf die betreffenden Abschnitte dieser Arbeit.

Ist die k-Sekanten-Hülle bezüglich der IK von k-Schmieg-Sekanten aus F für alle $k<k_0$ erklärt, und sind $S^{k_0}_{i_1}, S^{k_0}_{i_2} \dots S^{k_0}_{i_r}$ irgend welche k_0-Schmieg-Sekanten von F, ist ferner $S^{k_0}_{i_\varrho} \cap U^{n-k_0} = x^{k_0}_{i_\varrho}$ $(\varrho = 1, 2, \dots, r)$. $H(x^{k_0}_{i_1}, x^{k_0}_{i_2} \dots x^{k_0}_{i_r})$ die Punkthülle der $x^{k_0}_{i_\varrho}$ in $V(e^{k_0})$ und $H(S^{k_0-1}_{j_1}, S^{k_0-1}_{j_2} \dots S^{k_0-1}_{j_s})$ die (k_0-1)-Sekanten-Hülle bezüglich der IK der mit den $S^{k_0}_{i_1}, S^{k_0}_{i_2} \dots S^{k_0}_{i_r}$ inzidenten (k_0-1)-Schmieg-Sekanten von F, so wird unter der *k_0-Sekanten-Hülle* $H(S^{k_0}_{i_1}, S^{k_0}_{i_2} \dots S^{k_0}_{i_r})$ bezüglich der IK der $S^{k_0}_{i_1}, S^{k_0}_{i_2} \dots S^{k_0}_{i_r}$ verstanden: die Menge der k_0-Ebenen, die Punkte von $H(x^{k_0}_{i_1}, x^{k_0}_{i_2} \dots x^{k_0}_{i_r})$ mit (k_0-1)-Ebenen von $H(S^{k_0-1}_{j_1}, S^{k_0-1}_{j_2} \dots S^{k_0-1}_{j_s})$ verbinden.

Gleichbedeutend mit dieser Definition von $H(S^{k_0}_{i_1}, S^{k_0}_{i_2} \dots S^{k_0}_{i_r})$ ist folgende: Sind $S^u_{j_\lambda}$ $(\lambda = 1, 2, \dots, l_u)$ die mit den $S^k_{i_\varrho}$ $(\varrho = 1, 2, \dots, r)$ inzidenten u-Schmieg-Sekanten $(0 \leqq u \leqq k_0 - 1)$ und ist $S^u_{j_\lambda} \cap U^{n-u} = x^u_{j_\lambda}$, so besteht $H(S^{k_0}_{i_1}, S^{k_0}_{i_2} \dots S^{k_0}_{i_r})$ aus allen k_0-Ebenen, die einen Punkt von $H(S^0_{j_1}, S^0_{j_2}, \dots, S^0_{j_{l_0}})$ mit je einem Punkt aus $H(x^u_{j_1}, x^u_{j_2} \dots x^u_{j_{l_u}})$ verbinden $(1 \leqq u \leqq k_0)$.

Jede k-Ebene $\tilde{L}^k$ einer k-Sekanten-Hülle $(0 \leqq k \leqq n-1)$ bezüglich einer IK von k-Schmieg-Sekanten $S^k_{i_\sigma}$ $(\sigma = 1, 2, \dots, s)$ von F ist fremd zu U^{n-k-1}.

Dies ist richtig für $k=0$. Wir nehmen an, es sei richtig für $k=k_0-1$, und beweisen indirekt die Richtigkeit für $k=k_0$. $q \in U^{n-k_0-1}$ gehöre einer k_0-Ebene L^{k_0} von $H(S^{k_0}_{i_1}, S^{k_0}_{i_2} \dots S^{k_0}_{i_s})$ an, die also einen Punkt von $H(x^{k_0}_{i_1}, x^{k_0}_{i_2} \dots x^{k_0}_{i_s})$ mit einer (k_0-1)-Ebene L^{k_0-1} von $H(S^{k_0-1}_{j_1}, S^{k_0-1}_{j_2} \dots S^{k_0-1}_{j_t})$ verbindet, wobei die $S^{k_0-1}_{j_\tau}$ $(\tau = 1, 2, \dots, t)$ die mit den $S^{k_0}_{i_\sigma}$ $(\sigma = 1, 2, \dots, s)$ inzidenten (k_0-1)-Schmieg-Sekanten sind. Da $H(x^{k_0}_{i_1}, x^{k_0}_{i_2} \dots x^{k_0}_{i_s}) \subset V(e^{k_0})$ und $V(e^{k_0}) \cap U^{n-k_0-1} = \emptyset$ und $q \in \tilde{L}^{k_0} \cap U^{n-k_0-1}$, würde U^{n-k_0} von dieser Ebene $\tilde{L}^{k_0}$ mindestens in einer Geraden geschnitten werden, und es wäre $\tilde{L}^{k_0-1} \cap U^{n-k_0}$ nicht leer, im Widerspruch mit der Induktionsannahme

Ist L^r eine r-Ebene, die einer r-Sekanten-Hülle $H(S^r_{i_1} \dots S^r_{i_s})$ von F bezüglich einer IK angehört, so heiße L^r bezüglich U^{n-r-1} gleich orientiert wie die r-Schmieg-Sekanten $S^r_{i_1} \dots S^r_{i_s}$, wenn die Projektion eines positiven r-Simplexes von $S^r_{i_\sigma}$ $(\sigma = 1, 2, \dots, s)$ aus U^{n-r-1} in L^r ein positives r-Simplex in L^r ist. Dies gilt dann für alle positiven r-Simplexe in $S^r_{i_\sigma}$, deren Projektionen r-Simplexe sind und für alle σ (s. 5, vorletzter Absatz).

9. HILFSSATZ 3. VORAUSSETZUNG. $F = p_1, p_2 \dots$ sei eine unendliche Punktfolge, die den Forderungen 1 und 2 des Satzes 2 genügt. Es sei $S^{r-1}_{i,\varrho} = (p_i \dots p_{i+\varrho-1}, p_{i+\varrho+1} \dots p_{i+r})$ $(2 \leqq r \leqq n;\ 1 \leqq \varrho \leqq r-1)$, also $S^{r-1}_{i+\nu,\varrho-\nu} = (p_{i+\nu}, p_{i+\nu+1} \dots p_{i+\varrho-1}, p_{i+\varrho+1} \dots p_{i+\nu+r})$ $(1 \leqq \nu \leqq r-2;\ \nu < \varrho \leqq r-1)$.

BEHAUPTUNG. 1. *$S^{r-1}_{i,\varrho}$ gehört der $(r-1)$-Sekanten-Hülle $H(S^{r-1}_i\ S^{r-1}_{i+1})$ von S^{r-1}_i und S^{r-1}_{i+1} bezüglich jeder F zugeordneten IK an.*

2. *Das r-Simplex $\mathfrak{S}^r_{i,\varrho} = p_i \dots p_{i+\varrho-1}\, p_{i+\varrho+1} \dots p_{i+r+1}$ $(1 \leqq \varrho \leqq r;\ 1 \leqq r \leqq n)$ hat dieselbe Orientierung bezüglich U^{n-r-1} wie die r-Simplexe $\mathfrak{S}^r_i = p_i\, p_{i+1} \dots p_{i+r}$ und $\mathfrak{S}^r_{i+1} = p_{i+1}\, p_{i+2} \dots p_{i+r+1}$, ihre Projektionen aus U^{n-r-1} in S^r_i, S^r_{i+1} oder E^r gehören also derselben Klasse an.*

3. *Die Punktfolge $F_\varrho = p_1 \dots p_{i+\varrho-1}\, p_{i+\varrho+1} \dots$ erfüllt die Forderungen 1 und 2 des Satzes 2.*

VORBEMERKUNG. Behauptung 3. folgt aus 1. und 2.

Denn, daß die Forderung 1 erfüllt ist für F_ϱ, folgt direkt aus 2.; ferner haben die r-Schmieg-Sekanten von F_ϱ auch E^r zum Limes, da $S^r_{i,\varrho} \cap U^{n-r} \subset V(e^r)$ wegen $S^r_{i,\varrho} \in H(S^r_i, S^r_{i+1})$ gemäß Behauptung 1., und gemäß Behauptung 2. haben sie das orientierte E^r zum Limes.

9.1. Wir beweisen zuerst die Richtigkeit des Hilfssatzes für $r=2$ $(n \geqq 2)$.

9.1.1. *Betreff Behauptung* 1. Hier ist $\varrho=1$ und $S^r_{i,1}=(p_i, p_{i+2})$.

Es sei $S^1_i \cap U^{n-1} = X^1_i$. Da $x^2_i = S^2_i \cap U^{n-2}$, ist $x^2_i \in X^1_i$; da $x^1_i = S^1_i \cap U^{n-1}$, $x^1_{i+1} = S^1_{i+1} \cap U^{n-1}$, ist $x^1_i \in X^1_i$, $x^1_{i+1} \in X^1_i$. Da $x^1_i \in [p_i; p_{i+1})$ und $x^1_{i+1} \in [p_{i+1}; p_{i+2})$, liegt p_{i+1} zwischen p_i und x^1_i, p_{i+2} zwischen p_{i+1} und x^1_{i+1}.

Ferner sei $x^1_{i,1} = S^1_{i,1} \cap X^1_i$.

Von den drei Seiten des Dreiecks $p_{i+1}\, x^1_i\, x^1_{i+1}$ wird daher von $S^1_{i,1}$ die Seite $p_{i+1}\, x^1_{i+1}$ in p_{i+2}, aber nicht die Seite $p_{i+1}\, x^1_i$ geschnitten, also muß nach dem Axiom von PASCH auch die Seite $x^1_i\, x^1_{i+1}$ geschnitten werden. $x^1_{i,1}$ liegt daher zwischen x^1_i und x^1_{i+1}.

Die Gerade $S^1_{i,1}$ gehört daher der 1-Sekanten-Hülle $H(S^1_i, S^1_{i+1})$ bezüglich der IK an. Auf X^1_i folgen die Punkte x^1_i, $x^1_{i,1}$, x^1_{i+1} in dieser Reihenfolge aufeinander.

9.1.2. *Betreff Behauptung* 2. (für $r=2$). Ist p'_j die Projektion von p_j aus U^{n-3} in S^2_i, so erfüllt die Punktfolge $p_i, p_{i+1}, p_{i+2}, p'_{i+3} \ldots$ nach Hilfssatz 1 die Forderungen 1 und 2 des Satzes 2, und nach Hilfssatz 2 liegt p'_{i+3} oberhalb S^1_i in S^2_i. Das 2-Simplex $p_i\, p_{i+1}\, p'_{i+3}$ hat daher die gleiche Orientierung wie das 2-Simplex $p_i\, p_{i+1}\, p_{i+2}$; damit ist Behauptung 2 für $\varrho=2$ bewiesen.

Für $\varrho=1$ beweist man so: p'_{i+3} liegt oberhalb S^1_{i+1} in S^2_i. Es sei $S^{1\prime}_{i+2} = (p_{i+2}, p'_{i+3})$ und $S^{1\prime}_{i+2} \cap X^1_i = x^1_{i+2}$, $S^{1\prime}_{i+1,1} = (p_{i+1}, p'_{i+3})$ und $S^{1\prime}_{i+1,1} \cap X^1_i = x^1_{i+1,1}$.

Wie beim Beweis der Behauptung 1 ergibt sich, daß $x^1_{i+1,1}$ zwischen x^1_{i+1} und x^1_{i+2}, p'_{i+3} zwischen p_{i+2} und x^1_{i+2} liegt.

Aus 7. Zusatz, folgt, wenn man R^n, U^1, U^0 für $n=2$ ersetzt durch S^2_i, X^1_i, x^2_i, daß auf X^1_i die Punkte x^1_i, x^1_{i+1}, x^1_{i+2}, x^2_i in dieser Reihenfolge liegen, zusammen mit dem oben Bewiesenen erhält man auf X^1_i die Reihenfolge x^1_i, $x^1_{i,1}$, x^1_{i+1}, $x^1_{i+1,1}$, x^1_{i+2}, x^2_i.

Es ist das 2-Simplex $p_i\, p_{i+2}\, p_{i+1}$ und daher auch $p_i\, p_{i+2}\, x^1_i$ negativ, also $p_i\, p_{i+2}\, x^1_{i+2}$ und $p_i\, p_{i+2}\, p'_{i+3}$ positiv, w.z.z.w.

9.2. Wir nehmen Hilfssatz 3 als richtig an für $r=r_0-1 \geqq 2$ und beweisen ihn für $r=r_0>2$.

9.2.1. *Betreff Behauptung* 1. Es sei $Z^1_i = S^{r_0}_i \cap U^{n-r_0+1}$; $x^{r_0}_i \in Z^1_i$. $x^{r_0-1}_i$, $x^{r_0-1}_{i+1} \in Z^1_i$. Da $x^{r_0-1}_i \in [S^{r_0-2}_i; p_{i+r_0-1})$ und $x^{r_0-1}_{i+1} \in [S^{r_0-2}_{i+1}; p_{i+r_0})$, ist $p_i\, p_{i+1} \cdots p_{i+r_0-2}\, x^{r_0-1}_i$ bzw. $p_{i+1}\, p_{i+2} \cdots p_{i+r_0-1}\, x^{r_0-1}_{i+1}$ ein positives (r_0-1)-Simplex in $S^{r_0-1}_i$ bzw. in $S^{r_0-1}_{i+1}$.

Das (r_0-1)-Simplex $p_i \ldots p_{i+\varrho-1}\, p_{i+\varrho+1} \cdots p_{i+r_0-1}\, p_{i+\varrho}$ $(1 \leqq \varrho \leqq r_0-2)$ in $S^{r_0-1}_i$ ist positiv bzw. negativ, wenn $r_0-\varrho$ ungerade bzw. gerade ist. Da $S^{r_0-2}_{i,\varrho}$ nach Induktionsannahme der (r_0-2)-Sekanten-Hülle $H(S^{r_0-2}_i, S^{r_0-2}_{i+1})$ bezüglich der IK angehört und bezüglich $x^{r_0-1}_i = U^{n-r_0+1} \cap S^{r_0-1}_i$ gleich orientiert ist wie $S^{r_0-2}_i$, das (r_0-1)-Simplex $p_i \ldots p_{i+\varrho-1}\, p_{i+\varrho+1} \cdots p_{i+r_0-1}\, x^{r_0-1}_i$ also

positiv ist, *liegt* $p_{i+\varrho}$ *in* $S_i^{r_0-1}$ *auf der gleichen Seite von* $S_{i,\varrho}^{r_0-2}$ *wie* $x_i^{r_0-1}$ *bzw. auf entgegengesetzter Seite, wenn* $r_0-\varrho$ *ungerade bzw. gerade ist.* Das gilt auch noch für $\varrho=r_0-1$, denn dann liegt $p_{i+\varrho}=p_{i+r_0-1}$ auf derselben Seite von $S_i^{r_0-2}$ wie $x_i^{r_0-1}$ in $S_i^{r_0-1}$ gemäß Festsetzung von $x_i^{r_0-1}$.

Entsprechend erhält man für $S_{i+1}^{r_0-1}$ und $x_{i+1}^{r_0-1}$: $p_{i+\varrho}$ $(1<\varrho\leqq r_0-1)$ *liegt auf der gleichen Seite von* $S_{i+1,\varrho-1}^{r_0-2}$ *wie* $x_{i+1}^{r_0-1}$ *in* $S_{i+1}^{r_0-1}$ *bzw. auf der entgegengesetzten Seite, wenn* $r_0-\varrho$ *gerade bzw. ungerade ist.* Dies gilt auch für $\varrho=1$; denn dann liegt $p_{i+\varrho}=p_{i+1}$ auf derselben Seite von $S_{i+2}^{r_0-2}$ in $S_{i+1}^{r_0-1}$ wie $x_{i+1}^{r_0-1}$ bzw. auf entgegengesetzter, wenn $r_0-1=r_0-\varrho$ gerade bzw. ungerade ist.

Daraus folgt: *Liegen* $p_{i+\varrho}$ *und* $x_i^{r_0-1}$ *in* $S_i^{r_0-1}$ *auf der gleichen bzw. entgegengesetzten Seite von* $S_{i,\varrho}^{r_0-2}$, *so liegen* $p_{i+\varrho}$ *und* $x_{i+1}^{r_0-1}$ *in* $S_{i+1}^{r_0-1}$ *auf der entgegengesetzten bzw. gleichen Seite von* $S_{i+1,\varrho-1}^{r_0-2}$.

9.2.2. Die durch $x_i^{r_0-1}$, $x_{i+1}^{r_0-1}$ und $p_{i+\varrho}$ bestimmte 2-Ebene sei Z^2 $(1\leqq\varrho\leqq r_0-1)$. Es ist $Z^2<S_i^{r_0}$; $S_i^{r_0-1}$, $S_{i+1}^{r_0-1}$ und $S_{i,\varrho}^{r_0-1}$ werden daher von Z^2 in Geraden geschnitten, $S_i^{r_0-2}$, $S_{i+1}^{r_0-2}$, $S_{i,\varrho}^{r_0-2}$, $S_{i+1,\varrho-1}^{r_0-2}$ in Punkten, dabei ist $S_{i+1,0}^{r_0-2}=S_{i+2}^{r_0-2}$ und $S_{i,r_0-1}^{r_0-2}=S_i^{r_0-2}$.

Es sei $S_i^{r_0-1}\cap Z^2=(p_{i+\varrho},\ x_i^{r_0-1})=W_i^1$; $S_{i+1}^{r_0-1}\cap Z^2=(p_{i+\varrho},\ x_{i+1}^{r_0-1})=W_{i+1}^1$, $S_{i,\varrho}^{r_0-1}\cap Z^2=W_{i,\varrho}^1$; $W_i^1\cap W_{i+1}^1=p_{i+\varrho}$. $S_{i,\varrho}^{r_0-2}\cap Z^2=w_{i,\varrho}$; $S_{i+1,\varrho-1}^{r_0-2}\cap Z^2=w_{i+1,\varrho-1}$. Es ist $w_{i,\varrho}=S_i^{r_0-1}\cap S_{i,\varrho}^{r_0-1}\cap Z^2=W_i^1\cap W_{i,\varrho}^1$ und $w_{i+1,\varrho-1}=W_{i+1}^1\cap W_{i,\varrho}^1$; $W_{i,\varrho}^1=(w_{i,\varrho},\ w_{i+1,\varrho-1})$.

Das Ergebnis von 9.2.1. angewendet auf die Schnittgeraden und Schnittpunkte in Z^2 ergibt:

Liegen $p_{i+\varrho}$ und $x_i^{r_0-1}$ in W_i^1 auf der gleichen bzw. entgegengesetzten Seite von $w_{i,\varrho}$, so liegen $p_{i+\varrho}$ und $x_{i+1}^{r_0-1}$ in W_{i+1}^1 auf der entgegengesetzten bzw. gleichen Seite von $w_{i+1,\varrho-1}$, oder anders ausgedrückt: Liegt $w_{i,\varrho}$ nicht zwischen bzw. zwischen $p_{i+\varrho}$ und $x_i^{r_0-1}$, so liegt $w_{i+1,\varrho-1}$ zwischen bzw. nicht zwischen $p_{i+\varrho}$ und $x_{i+1}^{r_0-1}$, das bedeutet: $W_{i,\varrho}^1=(w_{i,\varrho},\ w_{i+1,\varrho-1})$ schneidet eine und nur eine der beiden Seiten $x_i^{r_0-1}\,p_{i+\varrho}$ und $x_i^{r_0-1}\,p_{i+\varrho}$ des Dreiecks $x_i^{r_0-1}\,x_{i+1}^{r_0-1}\,p_{i+\varrho}$, also wird nach dem Axiom von PASCH die Seite $x_i^{r_0-1}\,x_{i+1}^{r_0-1}$ von $W_{i,\varrho}^1$ in einem Punkt $x_{i,\varrho}$ geschnitten. Es folgen auf Z_i^1 die Punkte $x_i^{r_0-1}$, $x_{i,\varrho}$, $x_{i+1}^{r_0-1}$ in dieser Reihenfolge aufeinander.

$S_{i,\varrho}^{r_0-1}=(S_{i,\varrho}^{r_0-2},\ x_{i,\varrho})$ gehört demnach der (r_0-1)-Sekanten-Hülle $H(S_i^{r_0-1},\ S_{i+1}^{r_0-1})$ bezüglich der IK an, da nach Induktionsannahme $S_{i,\varrho}^{r_0-2}$ der (r_0-2)-Sekanten-Hülle $H(S_i^{r_0-2},\ S_{i+1}^{r_0-2})$ angehört.

9.2.3. *Betreff Behauptung 2 für* $r=r_0>2$.

Für $\varrho=r_0$ folgt die Richtigkeit aus Hilfssatz 2 und 1. Es sei nun $1\leqq\varrho\leqq r_0-1$.

Wir projizieren die Punktfolge F aus U^{n-r_0-1} in $S_i^{r_0}$. Es sei $(U^{n-r_0-1},\ p_\nu)\cap S_i^{r_0}=\overline{p}_\nu$ $(\nu=1,\ 2,\ \ldots)$ $\overline{p}_\nu=p_\nu$ für $i\leqq\nu\leqq i+r_0$ Die Punktfolge $\overline{p}_\nu$ erfüllt gemäß Hilfssatz 1 die Forderungen des Satzes 2

Es sei $\overline{S}_{i+\mu}^{m}=(\overline{p}_{i+\mu},\ \overline{p}_{i+\mu+1},\ \ldots,\ \overline{p}_{i+\mu+m})$ $(0\leqq m\leqq r_0-1;\ \mu=0,\ 1,\ \ldots)$

$\overline{S}_{i+1,\varrho-1}^{r_0-1}=(p_{i+1},\ \ldots,\ p_{i+\varrho-1},\ p_{i+\varrho+1},\ \ldots,\ \overline{p}_{i+r_0+1})$

$\overline{\mathfrak{S}}_{i+1,\varrho-1}^{r_0-1}=p_{i+1}\cdots p_{i+\varrho-1}\,p_{i+\varrho+1}\cdots\overline{p}_{i+r_0+1}$

$\overline{x}_{i+\mu}=\overline{S}_{i+\mu}^{r_0-1}\cap Z_i^1$; $\overline{x}_{i+1,\varrho-1}=\overline{S}_{i+1,\varrho-1}^{r_0-1}\cap Z_i^1$. $\overline{x}_i=x_i^{r_0-1}$; $\overline{x}_{i+1}=x_{i+1}^{r_0-1}$.

Es ist zu zeigen, daß das r-Simplex $\mathfrak{S}_{i,\varrho}^{r_0} = p_i \dots p_{i+\varrho-1}\, p_{i+\varrho+1} \dots p_{i+r_0}\, \bar{p}_{i+r_0+1}$ in $S_i^{r_0}$ positiv ist.

Wir wenden 7. Zusatz an. Dabei ist R^n durch $S_i^{r_0}$, U^k durch $U^{n-r_0+k} \cap S_i^{r_0}$ $(0 \leqq k \leqq r_0 - 1)$ zu ersetzen, also U^0 durch $U^{n-r_0} \cap S_i^{r_0} = x_i^{r_0}$.

Aus dem Zusatz folgt dann, daß auf Z_i^1 die Punkte $\bar{x}_i, \bar{x}_{i+1}\, \bar{x}_{i+2} \dots x_i^{r_0}$ in dieser Reihenfolge liegen und nach dem in 9.2.2 Bewiesenen liegt $x_{i,\varrho}$ bzw. $\bar{x}_{i+1,\varrho-1}$ zwischen $\bar{x}_i$ und $\bar{x}_{i+1}$ bzw. zwischen $\bar{x}_{i+1}$ und $\bar{x}_{i+2}$.

Es liegen also die Punkte $x_i^{r_0-1}$, $x_{i,\varrho}$, $x_{i+1}^{r_0-1}$, $\bar{x}_{i+1,\varrho-1}$, $\bar{x}_{i+2}$, $x_i^{r_0}$ in dieser Reihenfolge auf Z_i^1.

Es ist ferner außer

$$x_i^{r_0-1} \in [S_i^{r_0-2}, p_{i+r_0-1}); \quad x_{i+1}^{r_0-1} \in [S_{i+1}^{r_0-2}, p_{i+r_0}) \text{ auch } x_{i,\varrho} \in [S_{i,\varrho}^{r_0-2}, p_{i+r_0});$$
$$\bar{x}_{i+1,\varrho-1} \in [S_{i+1,\varrho-1}^{r_0-2}, \bar{p}_{i+r_0+1}).$$

Dabei wird berücksichtigt, daß nach Induktionsannahme $\mathfrak{S}_{i,\varrho}^{r_0-1}$, $\mathfrak{S}_{i+1}^{r_0-1}$ und $\mathfrak{S}_{i+1,\varrho-1}^{r_0-1}$ gleich orientiert sind und $S_{i,\varrho}^{r_0-2}$ und $S_{i+1,\varrho-1}^{r_0-2}$ der (r_0-2)-Sekanten-Hülle $H(S_i^{r_0-2}, S_{i+1}^{r_0-2}, S_{i+2}^{r_0-2})$ angehören.

Beachtet man ferner, daß $x_i^{r_0-1}$ und $p_{i+\varrho}$ in $S_i^{r_0-1}$ auf der gleichen oder entgegengesetzten Seite von $S_{i,\varrho}^{r_0-2}$ liegen, wenn $r_0 - \varrho$ ungerade oder gerade ist, und bedeutet $V(\mathfrak{S}_i^{r_0})$ das Vorzeichen des Simplex $\mathfrak{S}_i^{r_0}$, so erhält man:

$$\begin{aligned}
V(\mathfrak{S}_i^{r_0}) &= V(p_i\, p_{i+1} \dots p_{i+r_0-1}\, x_{i+1}^{r_0-1}) \\
&= (-1)^{r_0-\varrho-1}\, V(p_i\, p_{i+1} \dots p_{i+\varrho-1}\, p_{i+\varrho+1} \dots p_{i+r_0-1}\, p_{i+\varrho}\, x_{i+1}^{r_0-1}) \\
&= V(p_i\, p_{i+1} \dots p_{i+\varrho-1}\, p_{i+\varrho+1} \dots p_{i+r_0-1}\, x_i^{r_0-1}\, x_{i+1}^{r_0-1}) \\
&= V(p_i\, p_{i+1} \dots p_{i+\varrho-1}\, p_{i+\varrho+1} \dots p_{i+r_0-1}\, x_{i,\varrho}\, \bar{x}_{i+1,\varrho-1}) \\
&= V(p_i\, p_{i+1} \dots p_{i+\varrho-1}\, p_{i+\varrho+1} \dots p_{i+r_0}\, \bar{x}_{i+1,\varrho-1}) \\
&= V(p_i\, p_{i+1} \dots p_{i+\varrho-1}\, p_{i+\varrho+1} \dots p_{i+r_0}\, \bar{p}_{i+r_0+1}) \quad \text{w.z.z.w.}
\end{aligned}$$

10. Aus Hilfssatz 3, Behauptung 3 folgt nun, daß die Punktfolge $p_1, p_2 \dots$ auch nach Elimination eines und damit auch nach der Elimination von endlich vielen Punkten noch die Forderungen 1 und 2 des Satzes 2 erfüllt. Ist nun $1 \leqq \nu_0 < \nu_1 < \nu_2 < \dots < \nu_n$, so erfüllt die Punktfolge $p_1, p_2 \dots p_{\nu_0-1}, p_{\nu_0}, p_{\nu_1}, p_{\nu_2} \dots p_{\nu_n}, p_{\nu_n+1} \dots$ Forderung 1 von Satz 2, woraus folgt, daß das n-Simplex $p_{\nu_0}\, p_{\nu_1} \dots p_{\nu_n}$ positiv orientiert ist.

Damit ist Satz 2 bewiesen.

11. Es sei $F = p_1, p_2 \dots$ eine beschränkte monotone Punktfolge; die Koordinaten von p_i seien x_ν^i $(\nu = 1, 2, \dots, n)$ q sei ein beliebiger Punkt in R^n. $[q q_i)$ sei die zu $[p_{i-1} p_i)$ parallele Halbgerade.

Die Determinante $\begin{vmatrix} x_1^i & x_2^i & \dots & x_n^i & 1 \\ x_1^{i+1} & x_2^{i+1} & \dots & x_n^{i+1} & 1 \\ \dots & \dots & \dots & \dots & \dots \\ x_1^{i+n} & x_2^{i+n} & \dots & x_n^{i+n} & 1 \end{vmatrix}$ ist von 0 verschieden und hat für alle

$i \geqq 1$ das gleiche Vorzeichen. Dann ist auch die Determinante

$$\begin{vmatrix} x_1^{i+1} - x_1^i & x_2^{i+1} - x_2^i & \dots & x_n^{i+1} - x_n^i \\ x_1^{i+2} - x_1^{i+1} & x_2^{i+2} - x_2^{i+1} & \dots & x_n^{i+2} - x_n^{i+1} \\ \dots & \dots & \dots & \dots \\ x_1^{i+n} - x_1^{i+n-1} & x_2^{i+n} - x_2^{i+n-1} & \dots & x_n^{i+n} - x_n^{i+n-1} \end{vmatrix}$$

von 0 verschieden und hat immer das gleiche Vorzeichen.

Es sind daher je n aufeinanderfolgende Halbgerade $[q, q_{i+1})$, $[q, q_{i+2}) \dots [q, q_{i+n})$ und ihre Schnittpunkte $p'_{i+1}, p'_{i+2} \dots p'_{i+n}$ mit einer Hyperebene H^{n-1} für $i \geqq 1$ linear unabhängig und gleich orientiert; Die Punktfolge $F' = p'_2, p'_3 \dots$ erfüllt also Forderung 1 von Satz 2.

Ferner ist die k-Schmieg-Sekante S_i^k $(1 \leqq k \leqq n-1)$ parallel zur k-Ebene $T_i^k = (q, q_{i+1}, q_{i+2} \dots q_{i+n})$. Aus $\lim\limits_{i\to\infty} S_i^k = E^k$ (4) folgt daher $\lim\limits_{i\to\infty} T_i^k = F^k \| E^k$.

Der Halb-k-Schmieg-Sekante $[S_i^{k-1}, p_{i+k})$ entspricht die Halb-k-Ebene $[T_i^{k-1}, q_{i+k})$. Aus der Existenz einer Halb-k-Ebene als Limes von $[S_i^{k-1}, p_{i+k})$ folgt die Existenz einer Halb-k-Ebene F_+^k als Limes von $[T_i^{k-1}, q_{i+k})$ für $i \to \infty$.

Ist $S_i'^{k-1} = (p'_{i+1}, p'_{i+2} \dots p'_{i+k}) = T_i^k \cap H^{n-1}$, so ist $\lim\limits_{i\to\infty} [S_i'^{k-2}; p'_{i+k}) = \lim\limits_{i\to\infty} [T_i^{k-1}; q_{i+k}) \cap H^{n-1} = F_+^k \cap H^{n-1}$.

F' erfüllt daher Forderung 2' von Satz 2.

Von einem Punkt ausgehende Halbgerade bilden eine monotone Folge, wenn ihre Schnittpunkte mit einer Hyperebene eine monotone Folge bilden.

Man erhält also als Verallgemeinerung von Theorem 16 den

SATZ 3. *Die zu den gerichteten 1-Schmieg-Sekanten einer beschränkten monotonen Punktfolge $p_0, p_1 \dots$ im R^n parallelen von einem Punkt q ausgehenden Halbgeraden bzw. ihre Schnittpunkte p'_i $(i \geqq 1)$ mit einer zu q fremden Hyperebene bilden eine schließlich monotone Folge.*

Ist nun $\nu_1 < \nu_2 < \nu_3 < \cdots$, so ist auch $p_{\nu_1}, p_{\nu_2}, p_{\nu_3} \dots$ eine monotone Folge, da jede Teilfolge einer monotonen Folge auch eine monotone Folge ist. Nach Satz 3 bilden auch die von q ausgehenden zu $[p_{\nu_1}; p_{\nu_2})$, $[p_{\nu_2}; p_{\nu_3})$, $[p_{\nu_3}; p_{\nu_4})$ … parallelen Halbgeraden und deren unendlichen Teilfolgen schließlich monotone Folgen.

Daraus ergibt sich

SATZ 4. *Die von einem Punkt q ausgehenden zu den Halb-Sekanten $[p_{\nu_1}; p_{\nu_2})$, $[p_{\nu_3}; p_{\nu_4})$, $[p_{\nu_5}; p_{\nu_6})$ … $(\nu_1 < \nu_2 < \nu_3 < \cdots)$ einer beschränkten monotonen unendlichen Punktfolge $p_0, p_1, p_2 \dots$ parallelen Halbgeraden bilden eine schließlich monotone Halbgeradenfolge.*

Ist C ein monotoner Bogen, so ist jede im Sinn einer Orientierung von C auf C angeordnete Punktfolge monoton. Hieraus und aus Satz 4 folgt dann durch Grenzübergang von den Halbsekanten zu den Halbtangenten das am Anfang dieser Arbeit angeführte Theorem 28 von HJELMSLEV (s. HJELMSLEV, a.a.O., S. 64).

Erlangen, Hindenburgstr. 14

(Eingegangen am 22. September 1960)

ROSENBERG, A.
Math. Zeitschr. 76, 209–216 (1961)

Blocks and Centres of Group Algebras

By

ALEX ROSENBERG[1])

1. Introduction. In [3, 10B] BRAUER proved what is usually called the first main theorem on blocks. A somewhat different proof (although not quite complete, cf. Math. Rev. 17 (1956), p. 1182) was given by OSIMA in [4]. Both proofs use the theory of ordinary and modular characters in several essential points. Since blocks can be identified with the primitive idempotents of the centre of the group algebra while the characters correspond to the primitive idempotents of the algebra itself it seems desirable to supply a relatively simple proof which operates solely within the centres of the algebras in question and so does not use any tools from character theory. By combining the main ideas of [3] and [4] and recasting the definition of defect it is indeed possible to construct such a proof and this is carried out in the present note. It is hoped that the simplifications introduced will prove of some interest even though no really new results are obtained. In order to keep matters reasonably self-contained several lemmas with their proofs have been taken over verbatim from [3] and [4] and wherever this occurs appropriate references have been supplied.

In the present version no restrictions are placed on the ground field. Thus, it is not assumed, as it is in [3] and [4] that the field is large enough so that the absolutely irreducible representations of the group can be written in it. Finally, the main theorem and the methods of this paper are used to give a new proof of the theorem of BRAUER and NESBITT [2], which gives the number of blocks of highest defect.

This paper grew out of a series of talks given by the author in a seminar at the Institute for Advanced Study in 1955–56. Among the participants were I. REINER, P. ROQUETTE, E. STRAUSS, R. STEINBERG, and H. J. ZASSENHAUS. I express my thanks to all of them and especially to P. ROQUETTE for many illuminating comments. My thanks are also due to D. ZELINSKY for several suggestions.

2. Preliminaries and Notations. Let $\mathfrak{G}$ be a finite group of order g, p a fixed rational prime dividing g[2]) and $\mathfrak{F}$ a field of characteristic p. We shall write $\mathfrak{A} = \mathfrak{A}(\mathfrak{G}, \mathfrak{F})$ for the group algebra of $\mathfrak{G}$ over $\mathfrak{F}$ and let $\mathfrak{Z} = \mathfrak{Z}(\mathfrak{G}, \mathfrak{F})$ be the centre of $\mathfrak{A}$. If $\mathfrak{K}_1, \mathfrak{K}_2, \ldots, \mathfrak{K}_k$ are the classes of conjugate elements of $\mathfrak{G}$ and K_i, *the* i^{th} *class sum*, is the element of $\mathfrak{A}$ given by $\sum_{G \in \mathfrak{K}_i} G$, it is well

[1]) Written with the partial support of N.S.F. grants N.S.F.-G-1357 and N.S.F.-G-9508.

[2]) If p does not divide g, all the material remains formally valid but becomes trivial.

known that $K_1, K_2, \ldots, K_h$ are an $\mathfrak{F}$ basis of $\mathfrak{Z}$. In terms of this basis the multiplication in $\mathfrak{Z}$ is given by

$$K_i K_j = \sum_h a_{ijh} K_h \tag{2.1}$$

with a_{ijh} the residue class mod p of the number of ordered pairs (X, Y) of elements of $\mathfrak{G}$ satisfying

$$XY = G_h \qquad X \in \mathfrak{K}_i, \qquad Y \in \mathfrak{K}_j$$

for a fixed $G_h \in \mathfrak{K}_h$ [3, p. 409].

We shall use the notation $\mathfrak{N}(\mathfrak{S})$, $(\mathfrak{C}(\mathfrak{S}))$ for the normalizer (centralizer) of a subset, $\mathfrak{S}$, of $\mathfrak{G}$.

DEFINITION [3, (2.6)] [4, p. 176]. A p-Sylow subgroup of $\mathfrak{C}(G)$ is called a *defect group* of the class $\mathfrak{K}$ of G. Hence all defect groups of $\mathfrak{K}$ are conjugate and so have a common order, p^d. The integer d is called the *defect* of $\mathfrak{K}$.

We begin by introducing certain ideals in $\mathfrak{Z}$. In order to do this we need the following counting lemma which will be used several times.

LEMMA 2.1[3]. *Let $\mathfrak{D}$ be a p-subgroup of $\mathfrak{G}$, $\mathfrak{X}$ and $\mathfrak{Y}$ subsets of $\mathfrak{G}$ setwise invariant under conjugation by elements of $\mathfrak{D}$ and T a fixed element of $\mathfrak{C}(\mathfrak{D})$. Let S_T be the set of ordered pairs (X, Y) satisfying*

$$XY = T, \qquad X \in \mathfrak{X}, \qquad Y \in \mathfrak{Y}.$$

Then if $\mathfrak{Y} \cap \mathfrak{C}(\mathfrak{D})$ is empty, the cardinal of S_T is divisible by p.

PROOF. We partition S_T into non-overlapping classes by letting the class of a pair (X, Y) in S_T be all pairs (PXP^{-1}, PYP^{-1}) for all P in $\mathfrak{D}$. Hence the number of distinct elements in such a class is $[\mathfrak{D}:\mathfrak{D} \cap \mathfrak{C}(X) \cap \mathfrak{C}(Y)] = p^e$ for some integer e. Since $Y \notin \mathfrak{C}(\mathfrak{D})$ it follows that $\mathfrak{C}(Y)$ does not contain $\mathfrak{D}$ and so $e \neq 0$. Hence the cardinal of every class and thus of all of S_T is divisible by p.

DEFINITION. $\mathfrak{J}(\mathfrak{K})$ denotes the $\mathfrak{F}$-subspace of $\mathfrak{Z}$ spanned by the class sums of classes with defect groups conjugate to (not necessarily proper) subgroups of any defect group of $\mathfrak{K}$.

LEMMA 2.2 [4, Lemma 4]. *$\mathfrak{J}(\mathfrak{K})$ is an ideal of $\mathfrak{Z}$.*

PROOF. Let $K_j \in \mathfrak{J}(\mathfrak{K})$ and K_i be any class sum. Then $K_i K_j = \sum a_{ijh} K_h$. Lemma 2.1 with $\mathfrak{D} = \mathfrak{D}_h$ a defect group of $\mathfrak{K}_h$, $\mathfrak{X} = \mathfrak{K}_i$, $\mathfrak{Y} = \mathfrak{K}_j$, T a fixed element of $\mathfrak{K}_h$, together with the remark right after (2.1) show $a_{ijh} = 0$ whenever $\mathfrak{K}_j \cap \mathfrak{C}(\mathfrak{D}_h) = \emptyset$[4]. Hence if $a_{ijh} \neq 0$, $\mathfrak{D}_h$ will commute with some element in $\mathfrak{K}_j$ and so $\mathfrak{D}_h$ is conjugate to a subgroup of a defect group of $\mathfrak{K}_j$ and so of $\mathfrak{K}$. Hence $K_h \in \mathfrak{J}(\mathfrak{K})$.

3. Blocks. Let $E_1, E_2, \ldots$ be the primitive idempotents of $\mathfrak{Z}$. Since $\mathfrak{Z}$ is commutative they are uniquely determined by the conditions: (i) $1 = \sum E_i$,

[3] Lemma 2.1 is a slight generalization of [3, p. 426] and [4, Lemma 3].

[4] $\emptyset$ denotes the empty set.

(ii) $E_iE_j=0$ if $i\neq j$, and (iii) every E_i cannot be decomposed further into a sum of orthogonal nonzero idempotents. Moreover, $\mathfrak{Z}E_i$ is a completely primary algebra, i.e. if $\mathfrak{R}$ is the radical of $\mathfrak{Z}$, then $\mathfrak{Z}E_i/\mathfrak{R}E_i$ is a finite extension field of $\mathfrak{F}$. The canonical homomorphism, η_i, of $\mathfrak{Z}$ onto $\mathfrak{Z}E_i/\mathfrak{R}E_i$ has kernel the maximal ideal $\mathfrak{Z}(1-E_i)+\mathfrak{R}$. We shall refer to η_i as the irreducible representation of $\mathfrak{Z}$ associated with E_i. Conversely every $\mathfrak{F}$-algebra homomorphism of $\mathfrak{Z}$ into a finite extension field of $\mathfrak{F}$ arises from a primitive idempotent in this manner. Proofs of these well known results may be found for example in [*1*, chap. 9]. The E_i's will be called the *blocks* of $\mathfrak{Z}(\mathfrak{G},\mathfrak{F})$.

LEMMA 3.1. *Let E be a block, η the corresponding irreducible representation of $\mathfrak{Z}$, and $\mathfrak{J}$ an ideal of $\mathfrak{Z}$ such that $\eta(\mathfrak{J})\neq 0$. Then $E\in\mathfrak{J}$.*

PROOF. Since $\eta(\mathfrak{J})$ is a nonzero ideal of the field $\eta(\mathfrak{Z})$, we have $\eta(\mathfrak{J})=\eta(\mathfrak{Z})$. Consequently $\eta(E)\in\eta(\mathfrak{J})$. Thus there are elements $I\in\mathfrak{J}$, $R\in\mathfrak{R}$ and $Z\in\mathfrak{Z}$ with $I=E+Z(1-E)+R$. Hence $E=IE-RE$. Then, by raising both sides to the kth power, we find $E=I_k+(-1)^k(RE)^k$ with $I_k\in\mathfrak{J}$. Since RE is nilpotent the lemma is proved.

PROPOSITION 3.2. *Let $E=\sum f_iK_i$ be a block and η its corresponding irreducible representation. Then the index set $I=\{i\,|\,f_i\neq 0$ in $E=\sum f_iK_i\}$*[5]) *can be partitioned into two disjoint subsets A and B such that the defect groups of $\mathfrak{K}_\alpha$, $\alpha\in A$, are all conjugate, while the defect groups of $\mathfrak{K}_\beta$, $\beta\in B$, are conjugate to proper subgroups of the defect groups of the $\mathfrak{K}_\alpha$, $\alpha\in A$. Furthermore, if $\mathfrak{K}$ is any class whose defect groups are conjugate to proper subgroups of the defect groups of some $\mathfrak{K}_\alpha$, $\alpha\in A$, then $\eta(K)=0$, while for at least one $\alpha\in A$, we have $\eta(K_\alpha)\neq 0$.*

PROOF. Since $\eta(1)=\eta(E)\neq 0$ there is some $i_0\in I$ with $\eta(K_{i_0})\neq 0$. But then by Lemmas 2.2 and 3.1, $E\in\mathfrak{Z}(\mathfrak{K}_{i_0})$. Then if we let $A=\{i\in I\,|$ defect groups of $\mathfrak{K}_i$ are conjugate to those of $\mathfrak{K}_{i_0}\}$ and $B=I-A$, the first conclusion is clear. Again, if the defect groups of $\mathfrak{K}$ are conjugate to proper subgroups of the defect groups of $\mathfrak{K}_{i_0}$ and $\eta(K)\neq 0$, we would have $E\in\mathfrak{Z}(\mathfrak{K})$ which contradicts $f_{i_0}\neq 0$.

DEFINITION. Let $E=\sum\limits_{\alpha\in A}f_\alpha K_\alpha+\sum\limits_{\beta\in B}f_\beta K_\beta$, $f_\alpha, f_\beta\neq 0$, be a block written as in Proposition 3.2. Then the defect groups of $\mathfrak{K}_\alpha$, $\alpha\in A$, are called the *defect groups of E*. If their common order is p^d, d is called the *defect of E*.

REMARK. It follows from [*3*, 8A] or [*4*, Lemma 6 and Theorem 4] that this definition is equivalent to the one originally given by BRAUER.

The purpose of the main theorem of [*3*] is to exhibit a one-to-one correspondence of the blocks of $\mathfrak{Z}(\mathfrak{G},\mathfrak{F})$ whose defect groups are conjugate to $\mathfrak{D}$ with those of $\mathfrak{Z}(\mathfrak{N}(\mathfrak{D}),\mathfrak{F})$. To this end we make the following

DEFINITION [*3*, 7B]. Let $\mathfrak{D}$ be a p-subgroup of $\mathfrak{G}$. For any class sum K in $\mathfrak{Z}(\mathfrak{G},\mathfrak{F})$ let $\sigma(K)=$ the sum of all elements in $\mathfrak{K}\cap\mathfrak{C}(\mathfrak{D})$ if such elements exist, 0 otherwise.

[5]) By the use of modular characters it can be shown that all the $\mathfrak{K}_i$, $i\in I$, are p-regular classes [*4*, p. 177].

LEMMA 3.3 [3; 7B, 7C]. *If σ is extended to all of $\mathfrak{Z}(\mathfrak{G}, \mathfrak{F})$ by linearity, a homomorphism*

$$\sigma: \mathfrak{Z}(\mathfrak{G}, \mathfrak{F}) \to \mathfrak{Z}(\mathfrak{N}(\mathfrak{D}), \mathfrak{F})$$

is obtained. Ker σ *is spanned by the class sums K with $\mathfrak{K} \cap \mathfrak{C}(\mathfrak{D}) = \emptyset$.*

PROOF. Let $\mathfrak{H} = \mathfrak{N}(\mathfrak{D})$. Then clearly $\mathfrak{C}(\mathfrak{D})$ is normal in $\mathfrak{H}$ and so $\mathfrak{K} \cap \mathfrak{C}(\mathfrak{D})$ is closed under conjugation by elements of $\mathfrak{H}$ for any class $\mathfrak{K}$. Thus $\sigma(K)$ is a linear combination of class sums of $\mathfrak{Z}(\mathfrak{H}, \mathfrak{F})$ and hence σ is a linear mapping from $\mathfrak{Z}(\mathfrak{G}, \mathfrak{F})$ into $\mathfrak{Z}(\mathfrak{H}, \mathfrak{F})$.

It only remains to prove that $\sigma(K_i K_j) = \sigma(K_i)\sigma(K_j)$. Let $K = \sigma(K) + K'$ so that K' is the sum of the elements of $\mathfrak{K}$ not in $\mathfrak{C}(\mathfrak{D})$ and $\mathfrak{K}'$ is the set of summands of K'. By (2.1)

$$(3.1) \qquad (\sigma(K_i) + K_i')(\sigma(K_j) + K_j') = \sum a_{ijh}(\sigma(K_h) + K_h').$$

The elements of $\mathfrak{G}$ that occur as summands in $\sigma(K_i)\sigma(K_j)$ all lie in $\mathfrak{C}(\mathfrak{D})$ while those in $K_i'\sigma(K_j)$ and $\sigma(K_i)K_j'$ do not. Now applying Lemma 2.1 to $\mathfrak{X} = \mathfrak{K}_i'$, $\mathfrak{Y} = \mathfrak{K}_j'$ and T any element of $\mathfrak{K}_i'\mathfrak{K}_j' \cap \mathfrak{C}(\mathfrak{D})$ shows that the coefficient in $K_i'K_j'$ of any element of $\mathfrak{C}(\mathfrak{D})$ is 0. Thus (3.1) yields

$$\sigma(K_i)\sigma(K_j) = \sum a_{ijh}\sigma(K_h) = \sigma(K_i K_j).$$

LEMMA 3.4 [3, 10A]. *If $\mathfrak{K}$ is a class of $\mathfrak{G}$ having $\mathfrak{D}$ as one of its defect groups and $\mathfrak{H} = \mathfrak{N}(\mathfrak{D})$, then $\sigma(K)$ is a single class sum in $\mathfrak{Z}(\mathfrak{H}, \mathfrak{F})$ with defect group $\mathfrak{D}$. Conversely, if $\mathfrak{L}$ is a class of $\mathfrak{H}$ having $\mathfrak{D}$ for its (unique) defect group, then there is a class sum K in $\mathfrak{Z}(\mathfrak{G}, \mathfrak{F})$ with $\sigma(K) = L$.*

PROOF. There is an element $G \in \mathfrak{K}$ such that $\mathfrak{D}$ is a p-Sylow subgroup of $\mathfrak{C}(G)$. Hence $G \in \mathfrak{H}$. If $G' \in \mathfrak{K} \cap \mathfrak{C}(\mathfrak{D})$ there is an $X \in \mathfrak{G}$ such that $XGX^{-1} = G'$. Since $G' \in \mathfrak{C}(\mathfrak{D})$, $\mathfrak{D}$ is a p-Sylow subgroup of $\mathfrak{C}(G') = X\mathfrak{C}(G)X^{-1}$. But $X\mathfrak{D}X^{-1}$ is one also and so there is a $Y \in \mathfrak{C}(G')$ with $Y\mathfrak{D}Y^{-1} = X\mathfrak{D}X^{-1}$; hence $Y^{-1}X \in \mathfrak{H} = \mathfrak{N}(\mathfrak{D})$. Since Y commutes with G' we have $G' = Y^{-1}XGX^{-1}Y$ so that G' is conjugate to G in $\mathfrak{H}$. Finally, as we already saw in the proof of Lemma 3.3, if H is any element of $\mathfrak{H}$ conjugate to G in $\mathfrak{H}$, then $H \in \mathfrak{K} \cap \mathfrak{C}(\mathfrak{D})$. Thus $\sigma(K)$ is a class sum in $\mathfrak{Z}(\mathfrak{H}, \mathfrak{F})$.

For the converse, let $H \in \mathfrak{H}$ and suppose the centralizer, $\mathfrak{C}(H) \cap \mathfrak{H}$, of H in $\mathfrak{H}$ has $\mathfrak{D}$ as p-Sylow subgroup. Let $\mathfrak{D}_1 > \mathfrak{D}$ be a p-Sylow subgroup of $\mathfrak{C}(H)$. If $\mathfrak{D}_1 \neq \mathfrak{D}$, it is well known (e.g. [6; Chap. IV, Theorem 6]), that there is an element P_1 in $\mathfrak{N}(\mathfrak{D}) \cap \mathfrak{D}_1$ but not in $\mathfrak{D}$. But then the group generated by P_1 and $\mathfrak{D}$ would be a p-subgroup of $\mathfrak{C}(H) \cap \mathfrak{H}$ containing $\mathfrak{D}$ properly, which is impossible. Hence $\mathfrak{D}_1 = \mathfrak{D}$ is also a p-Sylow subgroup of $\mathfrak{C}(H)$. Thus if the class of H in $\mathfrak{G}$ is $\mathfrak{K}$ and its class in $\mathfrak{H}$ is $\mathfrak{L}$; the first part of the Lemma shows $\sigma(K) = L$.

PROPOSITION 3.5. *Let $\mathfrak{D}$ be a p-subgroup of $\mathfrak{G}$, let σ be the homomorphism of Lemma 3.3, $\mathfrak{H} = \mathfrak{N}(\mathfrak{D})$, and let $E = \sum_{\alpha \in A} f_\alpha K_\alpha + \sum_{\beta \in B} f_\beta K_\beta$ be a block written as in Proposition 3.2 for which $\mathfrak{D}$ is a defect group. Then $\sigma(E) = \sum_{\alpha \in A} f_\alpha \sigma(K_\alpha)$ is*

a nonzero idempotent of $\mathfrak{Z}(\mathfrak{H}, \mathfrak{F})$. *Furthermore* σ *induces a one-to-one mapping of the blocks of* $\mathfrak{Z}(\mathfrak{G}, \mathfrak{F})$ *into idempotents of* $\mathfrak{Z}(\mathfrak{H}, \mathfrak{F})$.

PROOF. By Proposition 3.2 the defect groups of $\mathfrak{K}_\beta$, $\beta \in B$, are conjugate to proper subgroups of $\mathfrak{D}$, hence $\mathfrak{K}_\beta \cap \mathfrak{C}(\mathfrak{D}) = \emptyset$ and $\sigma(K_\beta) = 0$. Since by Lemma 3.4 the $\sigma(K_\alpha)$ are distinct class sums and by Lemma 3.3 σ is an algebra homomorphism, $\sigma(E)$ is a nonzero idempotent. If E and E' are two blocks with $\mathfrak{D}$ as one of their defect groups and with $\sigma(E) = \sigma(E')$ then, writing them as in Proposition 3.2, we must have

$$E = \sum f_\alpha K_\alpha + \sum f_\beta K_\beta \qquad E' = \sum f_\alpha K_\alpha + \sum f'_\beta K_\beta,$$

where possibly some of the f_β, f'_β are 0. Now if η is the irreducible representation corresponding to E we have $\eta(K_\beta) = 0$. Hence $\eta(E) = \eta(E')$. But if E and E' really were distinct $\eta(E) = 1$ and $\eta(E') = 0$. Thus $E = E'$.

4. Normal p-subgroup. To a certain extent Proposition 3.5 has reduced the problem to the case where the defect group of the blocks in question is normal (and hence unique). Therefore throughout this section we shall assume that $\mathfrak{D}$ *is a normal* p*-subgroup of* $\mathfrak{G}$.

We first obtain a more precise description (Corollary 4.3) of $\mathfrak{J}(\mathfrak{K})$ for a class $\mathfrak{K}$ with defect group $\mathfrak{D}$.

LEMMA 4.1 [4, p. 183]. *The* $\mathfrak{F}$*-subspace,* $\mathfrak{V}$, *spanned by the class sums of classes with defect group* $\mathfrak{D}$ *is a subalgebra.*

PROOF. Let $K, K' \in \mathfrak{V}$. By Lemma 2.2, $KK' = \sum g_i K_i + \sum g_j K_j$ with $K_i \in \mathfrak{V}$ and K_j class sums of classes $\mathfrak{K}_j$ whose defect groups are conjugate to proper subgroups of $\mathfrak{D}$. The homomorphism σ defined above now is an algebra endomorphism of $\mathfrak{Z}(\mathfrak{G}, \mathfrak{F})$ and it is clear that if $K \in \mathfrak{V}$ then $\sigma(K) = K$ while $\sigma(K_j) = 0$. Hence $KK' = \sigma(K)\sigma(K') = \sigma(KK') = \sum g_i K_i$.

LEMMA 4.2. *Let* τ *be the algebra homomorphism* $\mathfrak{A}(\mathfrak{G}, \mathfrak{F}) \to \mathfrak{A}(\mathfrak{G}/\mathfrak{D}, \mathfrak{F})$ *induced by the natural group homomorphism* $\mathfrak{G} \to \mathfrak{G}/\mathfrak{D}$. *Then* Ker τ *is nilpotent. Furthermore, if* $\mathfrak{K} \cap \mathfrak{C}(\mathfrak{D}) = \emptyset$ *then* $\tau(K) = 0$ *and so* K *is nilpotent.*

PROOF. Since the characteristic of $\mathfrak{F}$ is p it is well known that $\mathfrak{A}(\mathfrak{D}, \mathfrak{F})$ is a local algebra whose radical (unique maximal ideal) is the ideal spanned by the elements $P - 1$[6]). Hence this ideal is nilpotent and so any sufficiently long product $\prod (P_i - 1)$, $P_i \in \mathfrak{D}$, is 0.

A routine calculation shows that Ker τ is the subspace of $\mathfrak{A}(\mathfrak{G}, \mathfrak{F})$ spanned by elements of the form $G(P - 1)$, $G \in \mathfrak{G}$, $P \in \mathfrak{D}$. Since $\mathfrak{D}$ is normal in $\mathfrak{G}$ we have $G(P-1)G'(P'-1) = GG'(P''-1)(P'-1)$ where P, P', P'' all are in $\mathfrak{D}$. Hence any sufficiently long product $\prod G_i(P_i - 1)$ is 0 and thus Ker τ is nilpotent.

Next, restricting τ to $\mathfrak{Z}(\mathfrak{G}, \mathfrak{F})$ we have for any class sum K, $\tau(K) = n\overline{K}$. Here $\overline{K}$ is a class sum in $\mathfrak{Z}(\mathfrak{G}/\mathfrak{D}, \mathfrak{F})$ and n is the number of distinct conjugates of a fixed element $G \in \mathfrak{K}$ which are congruent to G modulo $\mathfrak{D}$. Now suppose

[6]) This follows from [3, 3B]. An extremely simple direct proof is given in [5, p. 703].

$\mathfrak{K} \cap \mathfrak{C}(\mathfrak{D}) = \emptyset$ and set $\mathfrak{X} = G^{-1}\mathfrak{D}$, $\mathfrak{Y} = \mathfrak{K}$. Then n is simply the cardinal of the set S_1 as defined in Lemma 2.1, and this Lemma shows that $n \equiv 0 (p)$. Hence $\tau(K) = 0$, $K \in \operatorname{Ker} \tau$, and so is K nilpotent.

COROLLARY 4.3. *Let $\mathfrak{K}$ be a class with defect group $\mathfrak{D}$. Then $\mathfrak{Z}(\mathfrak{K}) = \mathfrak{V} \oplus \mathfrak{W}$*[7] *where $\mathfrak{V}$ is the subalgebra of $\mathfrak{Z}(\mathfrak{G}, \mathfrak{F})$ defined in Lemma 4.1 and $\mathfrak{W}$ is a nilpotent ideal of $\mathfrak{Z}(\mathfrak{G}, \mathfrak{F})$.*

PROOF. Let $\mathfrak{W}$ be the complement of $\mathfrak{V}$ spanned by the class sums with defect groups which are proper subgroups of $\mathfrak{D}$. By Lemma 3.3 $\mathfrak{W} = \operatorname{Ker} \sigma \cap \mathfrak{Z}(\mathfrak{K})$ and so is an ideal. Since $\mathfrak{W}$ admits a basis of nilpotent elements and $\mathfrak{Z}$ is commutative, $\mathfrak{W}$ is nilpotent.

PROPOSITION 4.4. *The defect groups of any block of $\mathfrak{Z}(\mathfrak{G}, \mathfrak{F})$ contain $\mathfrak{D}$. If the order of $\mathfrak{D}$ is p^d then any block of $\mathfrak{Z}(\mathfrak{G}, \mathfrak{F})$ of defect d lies in $\mathfrak{V}$.*

PROOF. Let E be a block with $\mathfrak{D}_E$ as one of its defect groups and suppose $E = \sum_{\alpha \in A} f_\alpha K_\alpha + \sum_{\beta \in B} f_\beta K_\beta$ in the notation of Proposition 3.2. If $\mathfrak{D} \not\subset \mathfrak{D}_E$ then for all $\alpha \in A$ and $\beta \in B$ we have $\mathfrak{K}_\alpha \cap \mathfrak{C}(\mathfrak{D}) = \mathfrak{K}_\beta \cap \mathfrak{C}(\mathfrak{D}) = \emptyset$. Thus $E \in \operatorname{Ker} \sigma$, a nilpotent ideal, which is impossible.

Now suppose E is a block of defect d. By the first part of the Proposition, the defect group of E is $\mathfrak{D}$. Hence by Corollary 4.3, $E = V + W$ with $V \in \mathfrak{V}$, $W \in \mathfrak{W}$. But there is an integer k with $W^{p^k} = 0$. Hence $E = V^{p^k} \in \mathfrak{V}$ by Lemma 4.1.

5. The main theorem. We now return to the general situation with $\mathfrak{D}$ a p-subgroup of $\mathfrak{G}$, $\mathfrak{H} = \mathfrak{N}(\mathfrak{D})$ and σ the homomorphism of $\mathfrak{Z}(\mathfrak{G}, \mathfrak{F})$ into $\mathfrak{Z}(\mathfrak{H}, \mathfrak{F})$ defined in 3. Only two steps remain in the proof of the main theorem.

LEMMA 5.1. *If E is a block of $\mathfrak{Z}(\mathfrak{G}, \mathfrak{F})$ with $\mathfrak{D}$ as one of its defect groups, then $\sigma(E)$ is a block of $\mathfrak{Z}(\mathfrak{H}, \mathfrak{F})$ with defect group $\mathfrak{D}$.*

PROOF. In view of Lemma 3.4 and Proposition 3.5 all that remains is to show that $\sigma(E)$ is primitive. If $\sigma(E)$ is not primitive, then $\sigma(E) = \sum_{i=1}^{t} F_i$, $t > 1$, where the F_i are orthogonal primitive idempotents of $\mathfrak{Z}(\mathfrak{H}, \mathfrak{F})$. Let η_i be the irreducible representation of $\mathfrak{Z}(\mathfrak{H}, \mathfrak{F})$ corresponding to F_i so that $\eta_i \mathfrak{Z}(\mathfrak{H}, \mathfrak{F})$ is a finite extension field of $\mathfrak{F}$. Since for each η_i we have $\eta_i \sigma(E) = 1$, the composite mappings $\eta_i \sigma$ are nonzero homomorphisms of $\mathfrak{Z}(\mathfrak{G}, \mathfrak{F})$ onto subfields, $\mathfrak{F}_i$, of $\eta_i \mathfrak{Z}(\mathfrak{H}, \mathfrak{F})$. These irreducible representations of $\mathfrak{Z}(\mathfrak{G}, \mathfrak{F})$ are all equivalent, i.e. there are $\mathfrak{F}$-isomorphisms φ_{ij} of $\mathfrak{F}_j$ onto $\mathfrak{F}_i$ such that $\varphi_{ij} \eta_j \sigma = \eta_i \sigma$.

By Proposition 4.4 the defect groups of each F_i contain $\mathfrak{D}$. If for some i_0 the defect groups of F_{i_0} were to contain $\mathfrak{D}$ properly, $\eta_{i_0}(\mathfrak{V}) = 0$ by Proposition 3.2. But then $\eta_{i_0} \sigma(E) = 0$, which is impossible. Thus all the F_i are in $\mathfrak{V}$. By Lemma 3.4 we may therefore write

$$F_i = \sum_j f_{ij} \sigma(K_j),$$

[7] Vector space direct sum.

where each $\mathfrak{K}_j$ has $\mathfrak{D}$ for one of its defect groups. Since $t>1$ we have

$$0=\eta_2(F_1)=\eta_2\sigma\left(\sum_j f_{1j}K_j\right)=\varphi_{21}\eta_1\sigma\left(\sum_j f_{1j}K_j\right)=\varphi_{21}\eta_1(F_1)=1,$$

which is a contradiction. Thus $\sigma(E)$ is primitive.

LEMMA 5.2. *Let the order of $\mathfrak{D}$ be p^d. Let F be a block of $\mathfrak{Z}(\mathfrak{H},\mathfrak{F})$ of defect d. Then there is a block, E, of $\mathfrak{Z}(\mathfrak{G},\mathfrak{F})$ having $\mathfrak{D}$ as one of its defect groups with $\sigma(E)=F$.*

PROOF. By Proposition 4.4, $F\in\mathfrak{V}$. Let η' be the irreducible representation of $\mathfrak{Z}(\mathfrak{H},\mathfrak{F})$ corresponding to F. Since by Lemma 3.4 $\mathfrak{V}$ is contained in $\sigma\mathfrak{Z}(\mathfrak{G},\mathfrak{F})$, we have $\eta'\sigma\neq 0$. Then $\eta'\sigma\neq 0$ is an irreducible representation of $\mathfrak{Z}(\mathfrak{G},\mathfrak{F})$ and so corresponds to a block E. Let $E=\sum_{\alpha\in A} f_\alpha K_\alpha+\sum_{\beta\in B} f_\beta K_\beta$ in the notation of Proposition 3.2. Then $\eta'\sigma(K_\beta)=0$ while $\eta'\sigma(K_{\alpha_0})\neq 0$ for some $\alpha_0\in A$. If $\mathfrak{K}_{\alpha_0}\cap\mathfrak{C}(\mathfrak{D})=\emptyset$ then $\sigma(K_{\alpha_0})=0$. Therefore the defect groups of $\mathfrak{K}_{\alpha_0}$, and so of E, contain subgroups conjugate to $\mathfrak{D}$. But if the defect groups of E were to contain $\mathfrak{D}$ properly, then by Proposition 3.2 and Lemma 3.4, $\eta'(\mathfrak{V})=0$ which contradicts $\eta'(F)=1$. Hence the defect groups of E are conjugate to $\mathfrak{D}$. Lemma 5.1 then shows that $\sigma(E)$ is primitive. But then since $\eta'\sigma(E)=\eta'(F)$, we have that $\sigma(E)=F$.

Combining Proposition 3.5, Lemmas 4.4, 5.1, and 5.2 yield

THEOREM 5.3. *Let $\mathfrak{G}$ be a finite group of order divisible by p, let $\mathfrak{D}$ be a subgroup of order p^d and let $\mathfrak{H}=\mathfrak{N}(\mathfrak{D})$ be its normalizer. Then if $\mathfrak{F}$ is any field of characteristic p there is a one-to-one correspondence between the blocks of $\mathfrak{Z}(\mathfrak{G},\mathfrak{F})$ with $\mathfrak{D}$ as one of their defect groups and the blocks of $\mathfrak{Z}(\mathfrak{H},\mathfrak{F})$ of defect d. The latter have $\mathfrak{D}$ as their defect group.*

If $\mathfrak{F}$ contains the gth (g = the order of $\mathfrak{G}$) roots of unity, Theorem 5.3 is [3, 10B] and [4, Theorem 6].

6. Blocks of highest defect. As an application of our methods we give a new proof of an old result of BRAUER and NESBITT [2, Theorem 2] and [3, 6D]. We recall that a class $\mathfrak{K}$ is said to be p-regular if the order of its elements is prime to p, and is called p-singular otherwise.

THEOREM 6.1. *Let the order of $\mathfrak{G}=p^a g'$, $(p,g')=1$. Let $\mathfrak{F}$ be an algebraically closed field of characteristic p. Then the number of blocks of defect a of $\mathfrak{Z}(\mathfrak{G},\mathfrak{F})$ is the number of p-regular classes of defect a.*

PROOF. A block of defect a has the p-Sylow subgroups for its defect groups. If $\mathfrak{D}$ is one of these and $\mathfrak{H}=\mathfrak{N}(\mathfrak{D})$ then Theorem 5.3 shows that the number of blocks of defect a of $\mathfrak{Z}(\mathfrak{G},\mathfrak{F})$ is the same as the number of blocks of $\mathfrak{Z}(\mathfrak{H},\mathfrak{F})$ of defect a. Since p^a is also the highest power of p dividing the order of $\mathfrak{H}$, Proposition 4.4 shows that these are all the blocks of $\mathfrak{Z}(\mathfrak{H},\mathfrak{F})$. Furthermore, Lemma 3.4 shows that the number of p-regular classes of defect a is the same in $\mathfrak{H}$ as in $\mathfrak{G}$, so that it is sufficient to work entirely in $\mathfrak{Z}(\mathfrak{H},\mathfrak{F})$.

As in Lemma 4.2, let τ be the homomorphism

$$\mathfrak{Z}(\mathfrak{H},\mathfrak{F})\to\mathfrak{Z}(\mathfrak{H}/\mathfrak{D},\mathfrak{F})$$

induced by the natural homomorphism $\mathfrak{H} \to \mathfrak{H}/\mathfrak{D}$. By Lemma 4.2, Ker τ is nilpotent. Hence it is a standard result that there is a one-to-one correspondence between the idempotents of $\mathfrak{Z}(\mathfrak{H}, \mathfrak{F})$ and those of Im τ (cf. e.g. [*1*, Theorem 9, 3 C]).

Now, since the order of $\mathfrak{H}/\mathfrak{D}$ is prime to p, $\mathfrak{A}(\mathfrak{H}/\mathfrak{D}, \mathfrak{F})$ is a semisimple algebra. Consequently, $\mathfrak{Z}(\mathfrak{H}/\mathfrak{D}, \mathfrak{F})$ and all its subalgebras are also semisimple. Since $\mathfrak{F}$ is algebraically closed this means that these algebras are just direct sums of copies of $\mathfrak{F}$. Therefore the number of blocks of defect a is $[\text{Im } \tau : \mathfrak{F}]$.

Next, let $\mathfrak{L}_i$, $i = 1, 2, \ldots, r$, be the p-regular classes of defect a of $\mathfrak{H}$. If $\mathfrak{L}$ is a class whose defect is less than a, Lemma 4.2 shows $\tau(L) = 0$. If $\mathfrak{L}$ is p-singular and $H \in \mathfrak{L}$, then $H = PR$ with P and R powers of H, the order of R prime to p, and the order of P a power of p[8]). But $\mathfrak{D}$ is the unique p-Sylow subgroup of $\mathfrak{H}$ and so $P \in \mathfrak{D}$ and $\tau(H) = \tau(R)$. Moreover $\mathfrak{C}(R) > \mathfrak{C}(H)$ and so R is in some $\mathfrak{L}_i$, $i = 1, 2, \ldots, r$. Hence Im τ is spanned by $\tau(L_i)$, $i = 1, 2, \ldots, r$.

Finally, let $H, H' \in \bigcup_{i=1}^{r} \mathfrak{L}_i$. If $\tau(H) = \tau(H')$ then $H = H'P$ with $P \in \mathfrak{D}$. But since $\mathfrak{D} < \mathfrak{C}(H')$ and the order of H is prime to p this means that $P = 1$. Hence τ induces a one-to-one mapping on $\bigcup_{1}^{r} \mathfrak{L}_i$. Thus each $\tau(L_i)$ is a single class sum of $\mathfrak{Z}(\mathfrak{H}/\mathfrak{D}, \mathfrak{F})$ and the $\tau(L_i)$, $i = 1, 2, \ldots, r$ are linearly independent elements of $\mathfrak{Z}(\mathfrak{H}/\mathfrak{D}, \mathfrak{F})$. Thus $[\text{Im } \tau : \mathfrak{F}] = r$ and Theorem 6.1 is proved.

REMARK. The hypothesis that $\mathfrak{F}$ be algebraically closed is too strong. All that is needed is to have $\mathfrak{F}$ large enough so that the absolutely irreducible representations of $\mathfrak{H}/\mathfrak{D}$ can be written in it: For then $\mathfrak{Z}(\mathfrak{H}/\mathfrak{D}, \mathfrak{F})$ is a direct sum of copies of $\mathfrak{F}$ and it can easily be seen that all its subalgebras are of the same form.

References

[*1*] ARTIN, E., C. NESBITT and R. M. THRALL: Rings with minimum condition. Ann Arbor 1948.

[*2*] BRAUER, R., and C. NESBITT: On the modular characters of groups. Ann. Math. **42**, 556–590 (1941).

[*3*] BRAUER, R.: Zur Darstellungstheorie der Gruppen endlicher Ordnung. Math. Z. **63**, 406–444 (1956).

[*4*] OSIMA, M.: Notes on blocks of group characters. Math. J. Okayama University **4**, 175–188 (1955).

[*5*] RIM, D. S.: Modules over finite groups. Ann. Math. **69**, 700–712 (1959).

[*6*] ZASSENHAUS, H. J.: The Theory of Groups. New York 1958.

[8]) If the order of $H = p^i m$, $(p, m) = 1$, and a, b are integers such that $1 = ap^i + bm$ then $P = H^{bm}$, $R = H^{ap^i}$.

Department of Mathematics, Northwestern University, Evanston, Ill. (U.S.A.)

(Eingegangen am 24. August 1960)

GROEMER, H.
Math. Zeitschr. 76, 217–225 (1961)

Lagerungs- und Überdeckungseigenschaften konvexer Bereiche mit gegebener Krümmung

Von

HELMUT GROEMER

1. Unter einem konvexen Bereich soll im folgenden stets eine beschränkte abgeschlossene konvexe ebene Punktmenge mit inneren Punkten verstanden werden. (K_i) $(i=1, 2, \ldots)$ sei eine Menge untereinander kongruenter konvexer Bereiche. Haben je zwei K_i höchstens Randpunkte gemeinsam, so heiße (K_i) eine Lagerung. Liegt hingegen jeder Punkt der Ebene in mindestens einem K_i, so heiße (K_i) eine Überdeckung. Falls eine Überdeckung vorliegt, soll außerdem stets vorausgesetzt werden, daß sich die Ränder von je zwei K_i in höchstens zwei Punkten schneiden. Diese Bedingung ist z. B. immer erfüllt, wenn alle K_i untereinander translationsgleich sind. Besitzt (K_i) eine Dichte d (für die Definition der Dichte siehe etwa FEJES TÓTH [1]), so gilt offenbar für eine Lagerung

$$d \leqq 1 \tag{1}$$

und für eine Überdeckung

$$d \geqq 1. \tag{2}$$

Im allgemeinen lassen sich (1) und (2) nicht verbessern. Der Zweck der folgenden Darlegungen ist, Verschärfungen von (1) und (2) anzugeben, wenn die obere oder die untere Grenze des Krümmungsradius des Randes von K_i gegeben ist.

2. Ist K ein beliebiger konvexer Bereich, so werde dessen Flächeninhalt mit $|K|$ und der Rand von K mit K^0 bezeichnet. Ohne es jedesmal hervorzuheben, möge stets vorausgesetzt werden, daß alle vorkommenden Dichten und Krümmungsradien existieren. Der Krümmungsradius ϱ von K^0 in einem Punkte $P \in K^0$ soll gemäß der folgenden bekannten Definition erklärt werden: Es sei X ein von P verschiedener Punkt auf K^0 und $\varrho(X, P)$ der Radius eines Kreises der durch X geht und K^0 in P berührt, das heißt, in P eine Tangente hat, die dort Stützgerade von K ist. Existiert

$$\varrho = \lim_{X \to P} \varrho(X, P), \tag{3}$$

so sei dies der Krümmungsradius von K^0 in P. Daß $\varrho = 0$ oder $\varrho = \infty$ ist, soll stets zulässig sein. Es sei noch darauf hingewiesen, daß man für alles Weitere auf die Existenz von (3) verzichten kann, wenn man in den Sätzen 1 und 3 als Krümmungsradius

$$\liminf_{X \to P} \varrho(X, P) \tag{4}$$

und in den Sätzen 2, 4 als Krümmungsradius

$$\overline{\lim_{X \to P}} \varrho(X, P) \tag{5}$$

definiert. Damit erhalten die angeführten Sätze für beliebige konvexe Bereiche einen Sinn.

3. Die folgenden Resultate sollen bewiesen werden:

SATZ 1. *Es sei K ein konvexer Bereich und (K_i) eine Lagerung oder Überdeckung mit zu K kongruenten K_i. Bedeutet d die Dichte von (K_i) und ist r die untere Grenze des Krümmungsradius des Randes von K, so gilt für eine Lagerung*

$$d \leqq \frac{1}{1 + (2\sqrt{3} - \pi)\frac{r^2}{|K|}} \tag{6}$$

und für eine Überdeckung

$$d \geqq \frac{1}{1 - \left(\pi - 3\frac{\sqrt{3}}{2}\right)\frac{r^2}{|K|}}. \tag{7}$$

Ist K ein Kreis vom Radius r, so gibt es Lagerungen und Überdeckungen, derart daß in (6) *bzw.* (7) *Gleichheit gilt. Für* (7) *ist der Kreis der einzige konvexe Bereich, für den Gleichheit gelten kann.*

SATZ 2. *Haben K und d die im Satz 1 genannte Bedeutung und ist R die obere Grenze des Krümmungsradius des Randes von K, so gilt für eine Lagerung*

$$d \leqq \frac{|K|}{6R^2 \operatorname{tg} \frac{|K|}{6R^2}} \tag{8}$$

und für eine Überdeckung

$$d \geqq \frac{|K|}{3R^2 \sin \frac{|K|}{3R^2}}. \tag{9}$$

In (8) *und in* (9) *kann Gleichheit dann und nur dann gelten, wenn K ein Kreis ist.*

Aus den Ungleichungen (6) und (8) oder noch einfacher aus den nachstehenden Ungleichungen (12) und (14) erhält man die beiden folgenden Verschärfungen des Minkowskischen Gitterpunktsatzes. Satz 3 wurde für stetig gekrümmte Bereiche auf anderem Wege schon von J. G. VAN DER CORPUT und H. DAVENPORT [2] bewiesen. Unter Gitterpunkt soll ein Punkt eines ebenen Gitters der Determinante 1 verstanden werden.

SATZ 3. *Ist K ein konvexer Bereich der bezüglich eines Gitterpunktes symmetrisch ist, und bedeutet r die untere Grenze des Krümmungsradius des Randes von K, so enthält K mindestens ein Gitterpunktpaar, wenn*

$$|K| \geqq 4 - (2\sqrt{3} - \pi)r^2 \tag{10}$$

ist.

SATZ 4. *Ist K wie in Satz 3 definiert und ist R die obere Grenze des Krümmungsradius des Randes von K, so enthält K mindestens ein Gitterpunktpaar, wenn*

$$|K| \geqq 6R^2 \operatorname{arc\,tg} \frac{2}{3R^2} \tag{11}$$

gilt.

Die Sätze 3 und 4 können nicht verschärft werden, in dem Sinne, daß für einen Kreis (und in (11) auch nur für einen Kreis) in (10) und (11) Gleichheit gilt und K außer dem Mittelpunkt keinen Gitterpunkt eines geeigneten Gitters der Determinante 1 im Innern enthält.

4. Nach bekannten Sätzen, für deren Beweis auf das Buch von FEJES TOTH [1] und die darin zitierte Literatur verwiesen werden kann, genügt es, die nachstehenden Ungleichungen (12), (13), (14), (15) zu beweisen. Dabei sollen, wie in allem weiteren, die folgenden Bezeichnungen verwendet werden. K ist ein nicht nur aus Randpunkten bestehender konvexer Bereich mit der oberen Grenze R und der unteren Grenze r des Krümmungsradius von K^0. $\overline{S}$ ist ein K umbeschriebenes und $\underline{S}$ ein K einbeschriebenes konvexes Sechseck. 3-, 4- und 5-Ecke sollen dabei stets als entartete Sechsecke mit einbezogen werden.

$$|\overline{S}| \geqq |K| + (2\sqrt{3} - \pi) r^2, \tag{12}$$

$$|\underline{S}| \leqq |K| - \left(\pi - \frac{3\sqrt{3}}{2}\right) r^2, \tag{13}$$

$$|\overline{S}| \geqq 6R^2 \operatorname{tg} \frac{|K|}{6R^2}, \tag{14}$$

$$|\underline{S}| \leqq 3R^2 \sin \frac{|K|}{3R^2}. \tag{15}$$

Man erkennt sofort, daß man zudem noch $r = 1$ bzw. $R = 1$ voraussetzen kann. Es reicht also hin für $r = 1$

$$|\overline{S}| \geqq |K| + (2\sqrt{3} - \pi), \tag{16}$$

$$|\underline{S}| \leqq |K| - \left(\pi - \frac{3\sqrt{3}}{2}\right) \tag{17}$$

und für $R = 1$

$$|\overline{S}| \geqq 6 \operatorname{tg} \frac{|K|}{6} \tag{18}$$

$$|\underline{S}| \leqq 3 \sin \frac{|K|}{3} \tag{19}$$

zu zeigen. Daß in diesen vier Ungleichungen und damit in (6), (7), (8), (9) Gleichheit gilt, wenn K ein Kreis ist und $\overline{S}$, $\underline{S}$ reguläre Sechsecke sind, erkennt man durch einfaches Nachrechnen. Daß umgekehrt aus dem Bestehen der Gleichheit in (7), (8) oder (9), wofür Gleichheit in (17), (18) bzw. (19) notwendig ist, folgt, daß K ein Kreis sein muß, wird aus dem Beweis dieser Ungleichungen folgen.

Der Beweis von (16), (17), (18), (19) beruht wesentlich auf den folgenden Hilfssatz, der wohl als bekannt vorausgesetzt werden darf.

HILFSSATZ 1. *Ist P ein Punkt von K^0, so liegt ein Kreis C' vom Radius r, der K^0 in P berührt, innerhalb K (d.h. $C' < K$), und ein Kreis C'' vom Radius R, der K^0 in P berührt außerhalb K (d.h. C'' und $K - K^0$ haben keine gemeinsamen Punkte).*

Die einzige Eigenschaft von r und R, die für alles Weitere benützt wird, ist die im Hilfssatz 1 genannte. Dies ist auch der Grund, warum man zur Definition von r und R die Ausdrücke (4) bzw. (5) anstatt (3) verwenden kann.

Für den Beweis von (18) wird sich der folgende Hilfssatz als nützlich erweisen. Bedeutet k irgendeine Kurve oder Strecke, so soll $|k|$ dessen Länge sein. Ist k eine Strecke, so sei k^∞ die Gerade die k enthält. Unter Kreis, Dreieck usw. sollen die entsprechenden Bereiche und nicht deren Rand verstanden werden.

HILFSSATZ 2. *Es sei C ein Einheitskreis und D ein Dreieck mit den folgenden Eigenschaften:*

a) Eine Seite g von D tangiert C.

b) Der g gegenüberliegende Eckpunkt von D ist in C.

c) Ist h die auf g gefällte Höhe von D, so ist $h \leqq 1$.

Wird $G = C \cap D$ gesetzt, dann gilt unter diesen Voraussetzungen

$$|D| \geqq \operatorname{tg} |G|. \tag{20}$$

BEWEIS. Bedeutet m das Minimum von $|D| - \operatorname{tg}|G|$ für alle Dreiecke D, die bei gegebenen h die Bedingungen a), b), c) erfüllen, so ist $m \geqq 0$ zu zeigen. Da $|D| - \operatorname{tg}|G|$ eine stetige Funktion der drei Eckpunkte von D ist und weil die Koordinaten dieser Eckpunkte beschränkt sind, wenn man D etwa die Bedingung $|D| - \operatorname{tg}|G| \leqq m + 1$ auferlegt, folgt aus dem bekannten Satz von Weierstraß, daß m für ein gewisses Dreieck D angenommen wird. Es kann daher vorausgesetzt werden, daß für D

$$m = |D| - \operatorname{tg}|G| \tag{21}$$

gilt. Es sei nun s der in D liegende Kreisbogen von C^0. s habe die Endpunkte P_1, P_2. Ist P der Berührungspunkt von s mit g, so sei s_1 der Teilbogen von s mit den Endpunkten P_1, P und s_2 der Teilbogen von s mit den Endpunkten P_2, P. Es soll zunächst

$$|s_1| = |s_2| \tag{22}$$

bewiesen werden. Angenommen, es wäre

$$|s_1| < |s_2|. \tag{23}$$

Es sei dann C' ein Einheitskreis, der g in einem Punkte P' berühre. Man definiere wieder $G' = D \cap C'$, so daß $(G \cup G') - (G \cap G')$ aus zwei Kreisbogendreiecken M_1, M_2 besteht (zwei Seiten dieser Dreiecke sind Einheitskreisbogen, die dritte eine Strecke). Der Rand von M_1 enthalte s_1, der von M_2

enthalte s_2. Wählt man P' auf der Seite von P auf der s_2 liegt, so erkennt man, daß

$$G' = (G \cup M_2) - M_1 \tag{24}$$

gilt. Läßt man den Abstand $d = \overline{PP'}$ gegen null streben, so ergibt sich, daß die beiden gekrümmten Seiten von M_1 die Längen s_1, $s_1 + o(1)$ und diejenigen von M_2 die Längen s_2, $s_2 + o(1)$ haben. Da der Winkel zwischen diesen Seiten für M_1 und M_2 derselbe ist, erhält man aus (23)

$$|M_2| > |M_1|,$$

womit aus (24)

$$|G'| > |G| \tag{25}$$

folgt. Bedeutet D' dasjenige Dreieck, das aus D durch die Bewegung entsteht, die C' in C und P' in P überführt, so erfüllt D' die Bedingungen a), b), c) und es ist wegen (25)

$$|D' \cap C| = |G'| > |G|.$$

Daraus ergäbe sich

$$m = |D| - \operatorname{tg}|G| > |D'| - \operatorname{tg}|D' \cap C|,$$

was der Definition von m widerspräche. Es gilt also (22). Nun läßt sich aber auch leicht zeigen, daß unter der Voraussetzung (22) angenommen werden darf, daß P der Mittelpunkt von g ist. Um dies einzusehen, betrachte man irgendein Dreieck D^+, dessen eine Seite g^+ in g^∞ liegt, mit C einen gemeinsamen Punkt hat und für die

$$|g^+| = |g| \tag{26}$$

gilt. Die anderen Seiten von D' sollen durch P_1 bzw. P_2 gehen. Aus (22) folgt, daß die Gerade, die P_1 und P_2 verbindet parallel zu g ist und daraus erhält man mit (26), wenn h^+ und G^+ analog wie h und g aber mittels D^+ anstatt D definiert werden

$$|h^+| = |h|,$$
$$|D^+| = |D|,$$
$$|G^+| = |G|.$$

Man erhält daraus, daß für jedes D^+

$$m = |D^+| - \operatorname{tg}|G^+|$$

ist. Da D^+ den Bedingungen a), b), c) genügt und weil es sicherlich ein D^+ gibt, für das g^+ durch P halbiert wird, folgt die Behauptung.

D und G sollen daher von nun an als symmetrisch bezüglich der durch P und den Mittelpunkt von C gehenden Geraden vorausgesetzt werden. Bezeichnet W den zu s gehörigen Kreissektor von C und setzt man $H = W - G$, so hat man wegen $h \leqq 1$

$$|D| \geqq \operatorname{tg}\frac{|s|}{2} - |H|$$

$$|G| = \frac{|s|}{2} - |H|$$

und daher

$$|D| - \operatorname{tg}|G| \geqq \operatorname{tg}\frac{|s|}{2} - |H| - \operatorname{tg}\left(\frac{|s|}{2} - |H|\right). \tag{27}$$

Durch Differenzieren erkennt man, daß die rechte Seite von (27) mit $|H|$ monoton ansteigend ist; das Minimum also für $|H|=0$ angenommen wird. Das ergibt

$$m = |D| - \operatorname{tg}|G| \geqq 0$$

und der Hilfssatz ist bewiesen. Soll in (18) Gleichheit gelten, so ist notwendig, daß $|H|=0$ also $h=1$ gilt.

Zum Beweis von (19) wird der folgende Hilfssatz benötigt.

HILFSSATZ 3. *Es sei C ein Einheitskreis und D ein Dreieck mit den folgenden Eigenschaften:*

a) Zwei Eckpunkte P_1, P_2 von D liegen auf C^0.

b) Der dritte Eckpunkt P_3 von D liegt bezüglich der durch P_1 und P_2 gehenden Geraden auf derselben Seite wie der Mittelpunkt von C.

c) Für die beiden Seiten l_1, l_2 von D, die P_3 und P_1 bzw. P_3 und P_2 verbinden gilt $l_1 \leqq 1$, $l_2 \leqq 1$.

Ist G der durch l_1, l_2 und einem Teilbogen von C^0 begrenzte konvexe Bereich, so gilt

$$|D| \leqq \tfrac{1}{2}\sin 2|G|. \tag{28}$$

BEWEIS. Zunächst erkennt man sofort, daß $l_1 = l_2 \leqq 1$ vorausgesetzt werden kann, weil man unter Erhaltung von $|D|$ und $|G|$ den Eckpunkt P_3 auf der durch P_3 gehenden Geraden, die zu der P_3 gegenüberliegenden Seite von D parallel ist, verschieben kann. Ist $s = G \cap C^0$ und W der durch s bestimmte Kreissektor, so gilt mit $H = W - G$

$$|D| = \tfrac{1}{2}\sin s - |H|$$
$$|G| = \tfrac{1}{2}\, s - |H|$$

und somit

$$|D| - \frac{1}{2}\sin\frac{s}{2} = \frac{1}{2}\sin s - |H| - \sin\left(\frac{s}{2} - |H|\right). \tag{29}$$

Der rechts stehende Ausdruck ist, wie man wieder durch Differenzieren erkennt, monoton fallend, nimmt also das Maximum für $|H|=0$ an. Daraus folgt (28). Gleichheit kann in (28), wie man unmittelbar sieht, nur dann gelten, wenn $l_1 = l_2 = 1$ ist.

5. Nun sollen die angeführten Sätze bewiesen werden.

BEWEIS VON SATZ 1. Zuerst soll (6) oder, was dafür hinreichend ist, (16) bewiesen werden. Es seien g_1 und g_2 zwei benachbarte Seiten von $\bar{S}$. Der Winkel zwischen g_1 und g_2 sei α_1. Derjenige Teilbereich von $\bar{S} - K$, dessen Rand in g_1, g_2 und K^0 liegt, heiße K_1. Man denke sich noch einen Einheitskreis, der g_1 und g_2 berührt, gegeben und nenne den Bereich, dessen Rand

in g_1, g_2 und der Einheitskreislinie liegt, C_1. Es soll zunächst

$$C_1 \subset K_1 \tag{30}$$

bewiesen werden. Wäre (30) nicht erfüllt, so gäbe es einen Punkt $P \in K_1^0$, der im Innern von C_1 liegt. Man könnte dann einen Einheitskreis C_1' konstruieren, der K_1^0 in P berührt. Da C_1 die Seiten g_1, g_2 berührt, muß ein Teil von C_1' außerhalb $\bar{S}$ liegen und nach Hilfssatz 1 kann dann K nicht innerhalb $\bar{S}$ sein, weil $C_1' \subset K$ angenommen werden darf. Aus (30) erhält man

$$|K_1| \geqq \operatorname{cotg} \frac{\alpha_1}{2} - \frac{1}{2}(\pi - \alpha_1) \tag{31}$$

und Entsprechendes gilt für die anderen Ecken mit den Winkeln α_i $(i = 1, \ldots, 6)$. Wegen $\sum_{i=1}^{6} \alpha_i = 4\pi$ folgt aus (31)

$$|\bar{S}| \geqq |K| + \sum_{i=1}^{5} \operatorname{cotg} \frac{\alpha_i}{2} - \operatorname{cotg}\left(\frac{1}{2}\sum_{i=1}^{5} \alpha_i\right) + \pi. \tag{32}$$

Sucht man nach den bekannten Methoden der Differentialrechnung das Minimum der rechten Seite von (32), so findet man, daß es für $\alpha_i = \frac{2\pi}{3}$ angenommen wird; man beachte dabei, daß $0 \leqq \frac{\alpha_i}{2} \leqq \frac{\pi}{2}$ sein muß. Dieser Wert für α_i in (32) eingesetzt ergibt (16).

Nun zum Beweis von (7), wofür es genügt (17) zu zeigen. g_1, g_2 seien zwei benachbarte Seiten von $\underline{S}$, die sich im Eckpunkt P von $\underline{S}$ treffen. Der Winkel zwischen g_1 und g_2 sei α_1. C sei ein Kreis vom Radius 1, der K^0 in P berührt. Die beiden Radien von C, die senkrecht auf g_1 bzw. g_2 stehen, sollen mit r_1, r_2 bezeichnet werden. Ist C_1 der Bereich, dessen Rand in g_1, g_2, r_1, r_2, C^0 liegt, und ist β der Winkel zwischen r_1 und der Verbindungsgeraden von P mit dem Mittelpunkt von C, so findet man

$$|C_1| = \tfrac{1}{2}(\pi - \alpha_1 - \sin\alpha_1 \cos(2\beta + \alpha_1)) \geqq \tfrac{1}{2}(\pi - \alpha_1 - \sin\alpha_1). \tag{33}$$

Aus Hilfssatz 1 folgt sofort, daß C_1 in K und daher in $K - \underline{S}$ enthalten ist. Für die zu den andern Ecken von $\underline{S}$ gehörigen C_i und α_i gilt dasselbe. Außerdem läßt sich zeigen, daß zwei verschiedene C_i keine inneren Punkte gemeinsam haben. Dies folgt daraus, daß C zur Gänze innerhalb K liegt, r_1 und r_2 daher die Seiten g_1, g_2 zwischen P und dem Mittelpunkt von g_1 bzw. g_2 schneiden. Aus (33) erhält man demnach mit $\sum_{i=1}^{6} \alpha_i = 4\pi$

$$|K - \underline{S}| \geqq \tfrac{1}{2}\left(2\pi - \sum_{i=1}^{5} \sin\alpha_i - \sin\left(4\pi - \sum_{i=1}^{5} \alpha_i\right)\right). \tag{34}$$

Das Minimum der rechten Seite von (34) wird wieder, unter Beachtung von $0 \leqq \alpha_i \leqq \pi$, für $\alpha_i' = \frac{2\pi}{3}$ angenommen, woraus sich sofort (17) ergibt. Soll in (17) Gleichheit gelten, so muß dies auch in (34) der Fall sein und daraus folgt, daß die C_i lückenlos aneinandergrenzen und $K = C$ ist.

BEWEIS VON SATZ 2. Zuerst soll (18) gezeigt werden, woraus dann (8) folgt. Aus Hilfssatz 1 erhält man, daß K in einem Kreis vom Radius 1 liegt. Es sei Q der Mittelpunkt dieses Kreises. Selbstverständlich darf angenommen werden, daß Q in K liegt. Durch Ziehen der Verbindungslinien der Eckpunkte von $\bar{S}$ mit Q wird $\bar{S}$ in sechs Dreiecke D_i ($i=1, \ldots, 6$) zerlegt, wobei zu bemerken ist, daß einige der Dreiecke zu einer Strecke degeneriert sein können. Die Q gegenüberliegende Seite g_i von D_i berühre K etwa in einem Punkt P_i. Nach Hilfssatz 1 gilt für einen Kreis C_i der g_i in P_i berührt mit den Bezeichnungen $K_i = D_i \cap K$, $G_i = D_i \cap C_i$

$$K_i \subset C_i. \tag{35}$$

Da der Abstand zwischen Q und P_i nicht größer als 1 ist, gilt dasselbe für die Höhe von D_i. Es sind demnach für jedes D_i und C_i die Bedingungen a), b), c) des Hilfssatzes 2 erfüllt und man hat zusammen mit (35)

$$|D_i| \geqq \operatorname{tg} |G_i| \geqq \operatorname{tg} |K_i|.$$

Daraus bekommt man

$$|\bar{S}| \geqq \sum_{i=1}^{6} \operatorname{tg} |K_i|. \tag{36}$$

Beachtet man jetzt, daß $\operatorname{tg} x$ für $0 \leqq x < \frac{\pi}{2}$ eine konvexe Funktion ist und daß wegen (35) $|K_i| < \frac{\pi}{2}$ gilt, so folgt (Jensensche Ungleichung)

$$\frac{1}{6} \sum_{i=1}^{6} \operatorname{tg} |K_i| \geqq \operatorname{tg} \left(\frac{1}{6} \sum_{i=1}^{6} |K_i| \right) = \operatorname{tg} \frac{|K|}{6}. \tag{37}$$

(36) und (37) ergeben (18). Gilt durchgehend das Gleichheitszeichen, so besteht der Rand von K, weil $K_i = G_i$ sein muß, aus sechs Einheitskreisbogen, deren Tangenten, wie aus dem am Ende des Beweises von Hilfssatz 2 Bemerkten hervorgeht, von dem Punkte Q den Abstand 1 haben. Daraus ergibt sich unmittelbar, daß K ein Kreis ist.

Nun ist noch (9) zu beweisen, wofür es genügt (19) zu zeigen. Q und D_i seien analog wie vorhin definiert. Die von Q ausgehenden Seiten der D_i sollen mit l_i bezeichnet werden. Für die l_i hat man

$$|l_i| \leqq 1. \tag{38}$$

K_i sei der durch je zwei benachbarte l_i ausgeschnittene konvexe Teilbereich von K, und G_i sei derjenige konvexe Bereich, der aus K_i dadurch entsteht, daß man $K_i - D_i$ durch ein passendes Einheitskreissegment ersetzt. Es soll nun

$$G_i \subset K_i \tag{39}$$

gezeigt werden. Nimmt man an, daß (39) nicht richtig sei, so gäbe es einen inneren Punkt P von G_i, der auf dem konvexen Begrenzungsbogen b_i von K_i ($b_i \subset K^0$) liegt. Konstruiert man den Einheitskreis C_i der b_i in P berührt, so muß offenbar einer der Eckpunkte von D_i, etwa P_i, nicht in C_i sein (P_i

bedeute einen von Q verschiedenen Eckpunkt von D_i). Andrerseits kann nach Hilfssatz 1 angenommen werden, daß b_i und daher auch P_i innerhalb C_i liegen. Gemäß der Definition von G_i und wegen (38) sind die Bedingungen a), b), c) des Hilfssatzes 3 erfüllt und man bekommt zusammen mit (39)

$$|\underline{S}| \leqq \tfrac{1}{2}\sum_{i=1}^{6} \sin 2|G_i| \leqq \tfrac{1}{2}\sum_{i=1}^{6} \sin 2|K_i|\,. \tag{40}$$

Weil $\sin x$ für $0 \leqq x < \pi$ konkav ist und $0 \leqq 2|K_i| < \pi$ ist, folgt wieder mittels der Jensenschen Ungleichung

$$\tfrac{1}{6}\sum_{i=1}^{6} \sin 2|K_i| \leqq \sin\left(\tfrac{1}{3}\sum_{i=1}^{6} |K_i|\right). \tag{41}$$

(40) und (41) ergeben (19). Der Zusatz über das Eintreten des Gleichheitszeichens folgt daraus, daß $G_i = K_i$ für Gleichheit notwendig ist und aus dem am Ende des Beweises von Hilfssatz 3 Gesagten.

Beweis der Sätze 3 und 4. Aus (12) und (10) oder (14) und (11) folgt, daß $|\bar{S}| \geqq 4$ ist, woraus die Behauptung leicht folgt. (Vgl. [2] oder K. Reinhardt [3]). Die Aussage bezüglich des Gleichheitszeichens ist ein bekanntes Resultat über die kritische Determinante des Kreises.

Literatur

[1] Fejes Toth, L.: Lagerungen in der Ebene, auf der Kugel und im Raum. Grundlehren der mathematischen Wissenschaften, Bd. 65. Berlin 1953.

[2] Corput, J. G. van der, and H. Davenport: On Minkowski's fundamental theorem in the geometry of numbers. Nederl. Akad. Wet., Proc. **49**, 701–707 (1946).

[3] Reinhardt, K.: Über die dichteste gitterförmige Lagerung kongruenter Bereiche in der Ebene und eine besondere Art konvexer Kurven. Abh. Math. Sem. Hansische Univ. **10**, 216–230 (1934).

Dept. of Math., Oregon State University, Corvallis, Oregon (U.S.A.)

(Eingegangen am 30. August 1960)

RIEGER, G. J.
Math. Zeitschr. 76, 226–234 (1961)

Über die Anzahl der Produktzerlegungen ganzer Zahlen

Von
G. J. RIEGER

Es sei $R(1)=T(1)=1$; für die natürliche Zahl $n>1$ bezeichne $R(n)$ bzw. $T(n)$ die Anzahl aller Darstellungen von n als Produkt von natürlichen Zahlen >1, wobei alle Darstellungen, die aus einer bestimmten durch Änderung der Reihenfolge hervorgehen, mit dieser als identisch bzw. von dieser als verschieden angesehen werden.

§ 1. Über $R(n)$

Man kennt den

SATZ (von SZEKERES und TURÁN). *Es ist*

$$\sum_{n\leq x} R(n) = \frac{x\, e^{2\sqrt{\log x}}}{2\sqrt{\pi}\log^{\frac{3}{4}} x}\left(1+O\left(\frac{1}{\sqrt{\log x}}\right)\right) \quad (x\to\infty)^{1}).$$

Wir wollen hier zeigen, daß sich für einen beliebig gegebenen Modul $k\geqq 1$ die Werte von $R(n)$ mit $(n,k)=1$ im Mittel gleichmäßig auf die $\varphi(k)$ primen Restklassen mod k verteilen (Satz 3).

Es bezeichne χ einen beliebigen Charakter mod k; $s=\sigma+it$; die Reihen

$$L(s,\chi) \doteq \sum_{n=1}^{\infty}\chi(n)\, n^{-s}\ ^{2})$$

und

$$\psi(s,\chi) \doteq \sum_{m=1}^{\infty}\frac{L(ms,\chi^m)-1}{m}$$

sind sicherlich konvergent für $\sigma>1$; $E(\chi)$ bedeute 1 für $\chi=\chi_0$ (Hauptcharakter) und 0 für $\chi\neq\chi_0$.

HILFSSATZ 1. *Für $\sigma>1$ ist*

$$\sum_{n=1}^{\infty}\chi(n)\, R(n)\, n^{-s} = e^{\psi(s,\chi)}.$$

BEWEIS. Für $\sigma>1$ ist einerseits

$$\begin{aligned} P(s,\chi) &\doteq \prod_{r=2}^{\infty}\left(1-\chi(r)\, r^{-s}\right)^{-1} \\ &= \prod_{r=2}^{\infty}\sum_{m=0}^{\infty}\chi(r^m)\, r^{-ms} \\ &= \sum_{n=1}^{\infty} R(n)\,\chi(n)\, n^{-s}, \end{aligned}$$

[1]) Wir beschränken uns auf zwei Glieder der in [7] gegebenen Entwicklung.
[2]) $\doteq$ bedeutet $=$ per definitionem.

für $\sigma>1$ ist andererseits

$$\log P(s,\chi) = -\sum_{r=2}^{\infty} \log\left(1-\chi(r)\,r^{-s}\right)$$

$$= \sum_{r=2}^{\infty}\sum_{m=1}^{\infty} \frac{\chi(r^m)}{m}\, r^{-ms}$$

$$= \sum_{m=1}^{\infty}\frac{1}{m}\sum_{r=2}^{\infty}\chi^m(r)\,r^{-ms}$$

$$= \sum_{m=1}^{\infty}\frac{1}{m}\left(L(ms,\chi^m)-1\right),$$

was zu zeigen war.

HILFSSATZ 2. *Für $\sigma>\frac{1}{2}$ ist*

$$g(s,\chi) \doteq \psi(s,\chi) - \frac{E(\chi)\,\varphi(k)}{k(s-1)} \tag{1}$$

regulär.

BEWEIS. Die Glieder der Reihe

$$g(s,\chi) = \frac{1}{1}\left(L(s,\chi) - \frac{E(\chi)\,\varphi(k)}{k(s-1)} - 1\right) + \frac{1}{2}\left(L(2s,\chi^2)-1\right)+\cdots$$

sind regulär für $\sigma>\frac{1}{2}$; wegen

$$\frac{1}{m}\left|L(ms,\chi^m)-1\right| < \sum_{n=2}^{\infty} n^{-m/2} \quad (m\geqq 4)$$

und wegen

$$\sum_{m=4}^{\infty}\sum_{n=2}^{N} n^{-m/2} < 2\left(\sum_{n=2}^{N} n^{-2} + \sum_{n=2}^{N} n^{-3}+\cdots\right)$$

$$= 2\sum_{n=2}^{N}\frac{1}{n(n-1)}$$

$$< 2$$

ist die Reihe selbst dort gleichmäßig konvergent, woraus die Behauptung folgt.

Eine unmittelbare Folge von Hilfssatz 2 ist

HILFSSATZ 3. *Für $r<\frac{1}{2}$ ist*

$$\frac{e^{g(1+re^{i\vartheta},\chi)}}{(1+re^{i\vartheta})^2} = e^{g(1,\chi)}\left(1+O(r)\right) \qquad (r\to 0).$$

HILFSSATZ 4. *Es ist*

$$\int_{-\pi/2}^{\pi/2} e^{-4y\sin^2\frac{\vartheta}{2}+i\vartheta}\,d\vartheta = \sqrt{\frac{\pi}{y}}\left(1+O\left(\frac{1}{y}\right)\right) \qquad (y\to\infty).$$

BEWEIS. Es ist

$$J \doteq \int_{-\pi/2}^{\pi/2} e^{-4y\sin^2\frac{\vartheta}{2}+i\vartheta}\,d\vartheta$$
$$= 2\int_0^{\pi/2} e^{-4y\sin^2\frac{\vartheta}{2}}\cos\vartheta\,d\vartheta$$
$$= \frac{2}{\sqrt{y}}\int_0^{\sqrt{2y}} e^{-u^2}\cdot\cos 2\arcsin\frac{u}{2\sqrt{y}}\cdot\left(1-\frac{u^2}{4y}\right)^{-\frac{1}{2}}du$$

vermöge

$$u = 2\sqrt{y}\sin\frac{\vartheta}{2}.$$

Wegen $0\leqq u\leqq\sqrt{2y}$ und daher $\frac{u^2}{4y}\leqq\frac{1}{2}$ kann man im Integranden nach Potenzen von $\frac{u^2}{4y}$ entwickeln und findet, wenn man noch

$$\int_{\sqrt{2y}}^{\infty} e^{-u^2}\left(1+O\left(\frac{u^2}{y}\right)\right)du = \int_{\sqrt{2y}}^{\infty} O(e^{-u})\,du$$
$$= O(e^{-\sqrt{2y}})$$

beachtet,

$$J = \frac{2}{\sqrt{y}}\int_0^{\sqrt{2y}} e^{-u^2}\left(1+O\left(\frac{u^2}{y}\right)\right)du$$
$$= \frac{2}{\sqrt{y}}\int_0^{\infty} e^{-u^2}\left(1+O\left(\frac{u^2}{y}\right)\right)du + O(e^{-\sqrt{2y}}),$$

woraus wegen

$$\int_0^{\infty} e^{-u^2}\,du = \frac{\sqrt{\pi}}{2}$$

die Behauptung folgt.

SATZ 1. *Es ist*

$$\sum_{n\leqq x}\chi(n)\,R(n)\log\frac{x}{n} = A_k\,x(\log x)^{-\frac{3}{4}}e^{2\sqrt{\frac{\varphi(k)\log x}{k}}}\left(E(\chi)+O\left(\frac{1}{\sqrt{\log x}}\right)\right)\quad(x\to\infty)$$

mit

$$A_k = \frac{1}{2\sqrt{\pi}}\left(\frac{\varphi(k)}{k}\right)^{\frac{1}{4}}e^{g(1,\chi_0)}$$

und mit einer nur von k abhängigen Konstanten im Restglied.

BEWEIS. Wegen Hilfssatz 1 ist nach einer geläufigen Formel[3])

$$\sum_{n\leqq x}\chi(n)\,R(n)\log\frac{x}{n} = \frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty}\frac{x^s}{s^2}\,e^{\psi(s,\chi)}\,ds.$$

[3]) Vgl. etwa [6], 268.

Auf Grund der (mit Hilfe der Weylschen Methode erzielbaren) Abschätzung

$$L(1+it,\chi)=O\left(\frac{\log t}{\log\log t}\right)\qquad(t\to\infty)\tag{2}$$

läßt sich unter Benutzung des Phragmén-Lindelöfschen Satzes der Integrationsweg von $\sigma=2$ nach $\sigma=1$ verlagern[4]), wobei der Punkt $s=1$ durch einen geeigneten Halbkreis auszuschließen ist, und man erhält

$$2\pi i\sum_{n\leqq x}\chi(n)\,R(n)\log\frac{x}{n}=J_1+J_2+J_3+J_4+J_5\qquad(x>e),$$

$$J_1\doteq i\int_{-\infty}^{-1}\frac{x^{1+it}}{(1+it)^2}\,e^{\psi(1+it,\chi)}\,dt,$$

$$J_2\doteq i\int_{-1}^{-\sqrt{\frac{\varphi(k)}{k\log x}}}\frac{x^{1+it}}{(1+it)^2}\,e^{\psi(1+it,\chi)}\,dt,$$

$$J_3\doteq\int_{\mathfrak{C}}\frac{x^s}{s^2}\,e^{\psi(s,\chi)}\,ds,$$

$$J_4\doteq i\int_{\sqrt{\frac{\varphi(k)}{k\log x}}}^{1}\frac{x^{1+it}}{(1+it)^2}\,e^{\psi(1+it,\chi)}\,dt,$$

$$J_5\doteq i\int_{1}^{\infty}\frac{x^{1+it}}{(1+it)^2}\,e^{\psi(1+it,\chi)}\,dt;$$

dabei bedeutet $\mathfrak{C}$ den Halbkreis

$$s=1+\sqrt{\frac{\varphi(k)}{k\log x}}\,e^{i\vartheta}\quad\left(-\frac{\pi}{2}\leqq\vartheta\leqq\frac{\pi}{2}\right).$$

Wegen Hilfssatz 2 ist

$$|J_2|=|J_4|=\left|x\int_{\sqrt{\frac{\varphi(k)}{k\log x}}}^{1}\frac{x^{it}}{(1+it)^2}\,e^{\frac{E(\chi)\varphi(k)}{kit}+g(1+it,\chi)}\,dt\right|=x\,O(1)\qquad(x\to\infty).$$

Wegen Hilfssatz 2 ist

$$|\psi(s,\chi)-L(s,\chi)|=O(1)\qquad(\sigma\geqq1;\ t\to\infty);$$

wegen (2) folgt daraus

$$|e^{\psi(1+it,\chi)}|=|e^{\psi(1+it,\chi)-L(1+it,\chi)}|\,|e^{L(1+it,\chi)}|=O(1)\,O\left(\sqrt{t}\right);$$

[4]) Vgl. auch [7], § 4.

daher ist

$$|J_1| = |J_5| < x \int_1^\infty \frac{1}{t^2} O(\sqrt{t})\, dt$$

$$= O(x) \qquad (x \to \infty).$$

Insgesamt haben wir damit gefunden:

$$\sum_{n \leqq x} \chi(n) R(n) \log \frac{x}{n} = \frac{J_3}{2\pi i} + O(x) \qquad (x \to \infty).$$

Im Falle $\chi = \chi_0$ ist wegen Hilfssatz 2

$$J_3 = i x \sqrt{\frac{\varphi(k)}{k \log x}} \int_{-\pi/2}^{\pi/2} e^{2\sqrt{\frac{\varphi(k)\log x}{k}} \cos\vartheta + i\vartheta} e^{g\left(1+\sqrt{\frac{\varphi(k)}{k\log x}} e^{i\vartheta}, \chi_0\right)} \left(1+\sqrt{\frac{\varphi(k)}{k\log x}} e^{i\vartheta}\right)^{-2} d\vartheta;$$

wegen $\cos\vartheta = 1 - 2\sin^2\vartheta/2$, Hilfssatz 3 und Hilfssatz 4 folgt daraus die Behauptung für $\chi = \chi_0$. Im Falle $\chi \neq \chi_0$ ist wegen Hilfssatz 2

$$J_3 = i x \sqrt{\frac{\varphi(k)}{k \log x}} \int_{-\pi/2}^{\pi/2} e^{\sqrt{\frac{\varphi(k)\log x}{k}} e^{i\vartheta}} e^{g\left(1+\sqrt{\frac{\varphi(k)}{k\log x}} e^{i\vartheta}, \chi\right)} \left(1+\sqrt{\frac{\varphi(k)}{k\log x}} e^{i\vartheta}\right)^{-2} d\vartheta$$

$$= i x \sqrt{\frac{\varphi(k)}{k \log x}} \int_{-\pi/2}^{\pi/2} O\left(e^{\sqrt{\frac{\varphi(k)\log x}{k}}}\right) O(1)\, d\vartheta;$$

daraus folgt die Behauptung für $\chi \neq \chi_0$. Damit ist Satz 1 bewiesen.

Mit Hilfe von

$$(3) \qquad \sum_{\chi} \chi(n) \bar{\chi}(l) = \begin{cases} \varphi(k) & \text{für } n \equiv l \bmod k, \\ 0 & \text{für } n \not\equiv l \bmod k \end{cases} \qquad ((l,k) = 1)$$

folgt aus Satz 1 sofort

SATZ 2. *Für* $(l, k) = 1$ *ist*

$$\sum_{\substack{n \leqq x \\ n \equiv l \bmod k}} R(n) \log \frac{x}{n} = B_k\, x\, (\log x)^{-\frac{1}{2}} e^{2\sqrt{\frac{\varphi(k)\log x}{k}}} \left(1 + O\left(\frac{1}{\sqrt{\log x}}\right)\right) \qquad (x \to \infty)$$

mit

$$B_k = \frac{A_k}{\varphi(k)}$$

und mit einer nur von k abhängigen Konstanten im Restglied.

Man verifiziert sofort

HILFSSATZ 5. *Für* $a_n \geqq 0$ $(n = 1, 2, \ldots)$ *und* $\lambda > 0$ *gilt*

$$\sum_{n \leqq x} a_n \log \frac{x}{n} - \sum_{n \leqq \frac{x}{1+\lambda}} a_n \log \frac{x}{n(1+\lambda)} \leqq \log(1+\lambda) \sum_{n \leqq x} a_n \leqq$$

$$\leqq \sum_{n \leqq x(1+\lambda)} a_n \log \frac{x(1+\lambda)}{n} - \sum_{n \leqq x} a_n \log \frac{x}{n}.$$

Die Anwendung von Hilfssatz 5 mit

$$a_n = \begin{cases} R(n) & \text{für } n \equiv l \bmod k, \\ 0 & \text{für } n \not\equiv l \bmod k \end{cases}$$

und

$$\lambda = \log^{-\frac{1}{2}} x$$

liefert wegen Satz 5 nach leichter Rechnung[5])

SATZ 3. *Für* $(l, k) = 1$ *gilt*

$$\sum_{\substack{n \leq x \\ n \equiv l \bmod k}} R(n) = B_k\, x (\log x)^{-\frac{1}{2}} e^{2\sqrt{\frac{\varphi(k) \log x}{k}}} \left(1 + O\left(\frac{1}{\sqrt{\log x}}\right)\right) \qquad (x \to \infty)$$

mit einer nur von k *abhängigen Konstanten im Restglied.*

Es ist nicht schwer, nach dem Vorbild von [7] weitere Glieder der Entwicklungen in den bewiesenen Sätzen anzugeben. Der Satz von SZEKERES und TURÁN ist der Spezialfall $k = 1$ von Satz 3.

§ 2. Über $T(n)$

Es sei P eine beliebige (endliche oder unendliche) Menge von nicht im Modul k aufgehenden Primzahlen[6]) mit den Eigenschaften:

(A) P enthält mindestens zwei Primzahlen p_1 und p_2,

(B) aus $\chi(p) = 1$ für alle $p \in P$ folgt $\chi = \chi_0$.

Die Menge Q enthalte außer 1 genau alle natürlichen Zahlen, in denen nur Primzahlen aus P aufgehen. Die Reihe

$$L(s, \chi_0, P) \doteq \sum_{n \in Q} n^{-s}$$

hat eine Konvergenzabszisse $\sigma_0(P)$ mit $0 \leqq \sigma_0(P) \leqq 1$. Die Funktionen

$$L(s, \chi, P) \doteq \sum_{n \in Q} \chi(n)\, n^{-s} \tag{4}$$

sind sicherlich für $\sigma > \sigma_0(P)$ erklärt und regulär, und es gilt dort

$$L(s, \chi, P) = \prod_{p \in P} (1 - \chi(p)\, p^{-s})^{-1}. \tag{5}$$

Die Eigenschaft (B) läßt sich auch so ausdrücken:

(B′) aus $L(s, \chi, P) = L(s, \chi_0, P)$ folgt $\chi = \chi_0$.

Denn: einerseits folgt aus $\chi(p) = 1$ für alle $p \in P$ sofort $L(s, \chi, P) = L(s, \chi_0, P)$ wegen (5); andererseits folgt aus $L(s, \chi, P) = L(s, \chi_0, P)$ zunächst $\Re L(\sigma, \chi, P) = L(\sigma, \chi_0, P)$ und daraus wegen

$$\chi(n)\, n^{-\sigma - it} \left\{ \begin{matrix} = \\ < \end{matrix} \right\} n^{-\sigma} \quad \text{für} \quad \begin{cases} \chi(n) = n^{it}, \\ \text{sonst} \end{cases} \tag{6}$$

mit $t = 0$ und wegen (4) sofort $\chi(n) = 1$ für alle $n \in Q$, insbesondere für alle $p \in P$.

[5]) Vgl. auch [7], 152–153.

[6]) In k aufgehende Primzahlen wären a ı f das folgende ohne Einfluß.

Im Falle $k>2$ ist (B) sicherlich dann nicht erfüllt, wenn P nur Primzahlen $\equiv 1 \bmod k$ enthält; (B) ist sicherlich dann erfüllt, wenn in P alle primen Restklassen mod k vertreten sind.

Für eine beliebige natürliche Zahl m bezeichne $T_m(n)$ die Anzahl aller Darstellungen der natürlichen Zahl $n>1$ als Produkt von m natürlichen Zahlen >1 mit Berücksichtigung der Reihenfolge; $T_0(n) \doteq 0$ für $n>1$; $T_m(1) \doteq 0$ für $m>0$, $T_0(1) \doteq 1$. Mit dieser Festsetzung gilt für $m \geqq 0$ offenbar

$$(7) \qquad (L(s,\chi,P)-1)^m = \sum_{n\in Q} \chi(n)\, T_m(n)\, n^{-s} \qquad (\sigma > \sigma_0(P)).$$

Über P setzen wir jetzt noch voraus:

es existiert eine reelle Zahl $\varrho \doteq \varrho(P) > \sigma_0(P)$ mit

$$(C) \qquad L(\sigma,\chi_0,P) \begin{cases} = 2 & \text{für } \sigma = \varrho(P), \\ < 2 & \text{für } \sigma > \varrho(P). \end{cases}$$

Die Eigenschaft (C) ist sicherlich dann erfüllt, wenn P nur endlich viele Primzahlen oder alle nicht in k aufgehenden Primzahlen enthält. Für $\sigma > \varrho(P)$ ist

$$|L(s,\chi,P)-1| \leqq L(\sigma,\chi_0,P)-1 < 1.$$

Summieren wir daher (7) über m, so folgt wegen

$$\sum_{m=0}^{\infty} T_m(n) = T(n)$$

sofort

$$f(s,\chi,P) \doteq \frac{1}{2-L(s,\chi,P)} = \sum_{n\in Q} \chi(n)\, T(n)\, n^{-s} \qquad (\sigma > \varrho(P)).$$

Wegen (3) folgt daraus

HILFSSATZ 6. *Für $(l,k)=1$ und $\sigma > \varrho(P)$ gilt*

$$(8) \qquad \sum_{\chi} \bar{\chi}(l)\, f(s,\chi,P) = \varphi(k) \sum_{\substack{n\in Q \\ n\equiv l \bmod k}} T(n)\, n^{-s}.$$

Die Funktionen $f(s,\chi,P)$ sind regulär für $\sigma > \varrho(P)$; für $\sigma = \varrho(P)$ gilt

HILFSSATZ 7. *Die Funktionen $f(\varrho+it,\chi,P)$ $(\chi \neq \chi_0)$ sind stetige Funktionen von t; $f(\varrho+it,\chi_0,P)$ ist bis auf den einfachen Pol für $t=0$ mit dem Residuum $-(L'(\varrho,\chi_0,P))^{-1}$ eine stetige Funktion von t*[7].

BEWEIS. Es ist die Gleichung

$$L(\varrho+it,\chi,P) = 2$$

oder, was dafür notwendig ist,

$$(9) \qquad \Re\, L(\varrho+it,\chi,P) = L(\varrho,\chi_0,P)$$

[7]) Offenbar ist $L'(\varrho(P),\chi_0,P) < 0$.

zu untersuchen. Nun ist

$$\chi(n) = e^{\frac{2\pi i a}{\varphi(k)}} \tag{10}$$

mit ganzrationalem $a = a(n, \chi)$. Wegen (6) mit $\sigma = \varrho$ folgen aus (9) bei festem χ für t die unendlich vielen simultanen Bedingungsgleichungen

$$\chi(n) = n^{it}$$

für alle $n \in Q$ oder wegen (10)

$$t = \frac{\alpha}{\log n} \qquad (\alpha \text{ rational}),$$

also wegen (A) insbesondere

$$t = \frac{\alpha_1}{\log p_1} = \frac{\alpha_2}{\log p_2} \qquad (\alpha_1, \alpha_2 \text{ rational}),$$

was nur für $t = 0$ möglich ist. Für $t = 0$, $\chi \neq \chi_0$ ist aber (9) unmöglich wegen (B) und (6). Die Funktion $f(\varrho + it, \chi, P)$ ist also höchstens unstetig für $t = 0$, $\chi = \chi_0$; in diesem Falle liegt offenbar ein einfacher Pol vor mit dem behaupteten Residuum, womit Hilfssatz 7 bewiesen ist.

Wegen Hilfssatz 7 ist der Taubersatz von WIENER und IKEHARA auf (9) anwendbar und liefert

SATZ 4. *Unter den Voraussetzungen* (A), (B); (C) *über* P *gilt für* $(l, k) = 1$ *die Beziehung*

$$\sum_{\substack{n \equiv l \bmod k \\ n \in Q,\, n \leq x}} T(n) = - \frac{(1 + o(1))\, x^{\varrho(P)}}{\varphi(k)\, \varrho(P)\, L'(\varrho(P), \chi_0, P)} \qquad (x \to \infty).$$

Die Werte von $T(n)$ mit $n \in Q$ verteilen sich also im Mittel gleichmäßig auf die $\varphi(k)$ primen Restklassen mod k. Für $k = 1$ liefert Satz 3 einen Satz von HILLE[8]; im Spezialfall $k = 1$, $P = P_0$ (Menge *aller* Primzahlen) kennt man die Ergebnisse von KALMÁR[9] und IKEHARA[10] mit verschiedenen Restgliedern.

Satz 4 kann auf arithmetische Halbgruppen verallgemeinert werden. Wir begnügen uns jedoch mit der Menge aller ganzen Ideale eines algebraischen Zahlkörpers K. Dazu sei $\mathfrak{f}$ ein beliebiges ganzes Ideal von K; die zur Einteilung aller zu $\mathfrak{f}$ teilerfremden Ideale in (engste) Idealklassen $\mathfrak{H} \bmod \mathfrak{f}$ gehörigen Charaktere bezeichnen wir mit X, insbesondere mit X_0 den Hauptcharakter; $h(\mathfrak{f})$ bedeute die Klassenzahl. Es sei $\mathfrak{P}$ eine beliebige (endliche oder unendliche) Menge von nicht in $\mathfrak{f}$ aufgehenden Primidealen mit den entsprechenden Eigenschaften (A)[11], (B), (C); die Menge $\mathfrak{Q}$ enthalte außer dem Einheitsideal $\mathfrak{o}$ genau alle ganzen Ideale von K, in denen nur Primideale von $\mathfrak{P}$ aufgehen; dabei ist

$$L(s, \mathsf{X}, \mathfrak{P}) \doteq \sum_{\mathfrak{a} \in \mathfrak{Q}} \mathsf{X}(\mathfrak{a})\, N\mathfrak{a}^{-s}$$

---

[8]) Vgl. neben [3] auch [1] und [2], insbesondere das Addendum von HILLE in [2].
[9]) Vgl. [5].
[10]) Vgl. [4].
[11]) $\mathfrak{P}$ enthält mindestens zwei Primideale $\mathfrak{p}_1$ und $\mathfrak{p}_2$ mit $N\mathfrak{p}_1 \neq N\mathfrak{p}_2$.

für $\sigma > \sigma_0(\mathfrak{P})$, wo $N\mathfrak{a}$ die Norm des Ideals $\mathfrak{a}$ und $\sigma_0(\mathfrak{P})$ die Konvergenzabszisse von $L(s, \chi_0, \mathfrak{P})$ bedeutet; es ist $0 \leqq \sigma_0(\mathfrak{P}) \leqq 1$. Für $\mathfrak{a} \neq \mathfrak{o}$ bezeichne $T(\mathfrak{a})$ die Anzahl aller Darstellungen von $\mathfrak{a}$ als Produkt von ganzen Idealen $\neq \mathfrak{o}$ mit Berücksichtigung der Reihenfolge der Faktoren; $T(\mathfrak{o}) \doteq 1$. Der Beweis von Satz 4 überträgt sich auf den vorliegenden Fall und liefert

SATZ 5. *Unter den Voraussetzungen* (A), (B), (C) *über* $\mathfrak{P}$ *gilt in jedem algebraischen Zahlkörper* K *die Beziehung*

$$\sum_{\substack{\mathfrak{a} \in \mathfrak{H} \bmod f \\ \mathfrak{a} \in \mathfrak{Q}, N\mathfrak{a} \leqq x}} T(\mathfrak{a}) = -\frac{(1 + o(1))\, x^{\varrho(\mathfrak{P})}}{h(f)\, \varrho(\mathfrak{P})\, L'(\varrho(\mathfrak{P}), \chi_0, \mathfrak{P})} \qquad (x \to \infty).$$

Die Abhängigkeit des Fehlergliedes vom Modul in den Sätzen soll nicht weiter untersucht werden.

Zu einer interessanten Verallgemeinerung der behandelten Fragestellungen gelangt man, wenn man an Stelle von $R(n)$ und $T(n)$ die Funktionen $R(n, y)$ und $T(n, y)$ verwendet; dabei bezeichne $R(n, y)$ bzw. $T(n, y)$ die Anzahl der Darstellungen der natürlichen Zahl $n > 1$ als Produkt von natürlichen Zahlen a mit $1 < a \leqq y$, wobei alle Darstellungen, die aus einer bestimmten durch Änderung der Reihenfolge hervorgehen, mit dieser als identisch bzw. von dieser als verschieden angesehen werden; $y = y(n)$ ist hier eine geeignete positivwertige Funktion von n. Auch darauf gehen wir hier nicht ein.

Literatur

[1] ERDÖS, P.: On some asymptotic formulas in the theory of the "factorisatio numerorum". Ann. Math. (2) **42**, 989–993 (1941).

[2] ERDÖS, P.: Corrections to two of my papers. Ann. Math. (2) **44**, 647–651 (1943).

[3] HILLE, E.: A problem in "factorisatio numerorum". Acta Arithm. **2**, 134–144 (1937).

[4] IKEHARA, S.: On KALMÁR's problem in „factorisatio numerorum". Proc. phys.-math. Soc. Jap. (3) **21**, 208–219 (1939).

[5] KALMÁR, L.: Über die mittlere Anzahl der Produktdarstellungen der Zahlen. 1. Mitteilung. Acta Litt. Sci. Szeged **5**, 95–107 (1930–1932).

[6] LANDAU, E.: Handbuch der Lehre von der Verteilung der Primzahlen. New York 1953.

[7] SZEKERES, G., u. P. TURÁN: Über das zweite Hauptproblem der "factorisatio numerorum". Acta Litt. Sci. Szeged **6**, 143–154 (1932–1934).

Department of Mathematics and Statistics, Purdue University, Lafayette, Indiana, USA.

(Eingegangen am 30. September 1960)

HEATON, R.
Math. Zeitschr. 76, 235–239 (1961)

A Note on Polynomial Cocycles

By
ROBERT HEATON

1. INTRODUCTION. Polynomial cocycles and coboundaries over a field K are defined in [4]. The definition is analogous to that given by EILENBERG and MACLANE in the case of abstract groups, see [1], where the variables in the coboundary formula are taken instead to be indeterminates. As expected a polynomial coboundary is a cocycle. In fact, when the characteristic of K is zero, it is shown in [4] that the only polynomial n-cocycles with $n>1$ are coboundaries. On the other hand, if K is of characteristic p, the situation is much more complicated, and for any positive integer n there are polynomial n-cocycles which are not coboundaries. The complete structure of polynomial 1, 2, and 3-cocycles for characteristic p, is given in [3] and [4], but the general dimensional problem remains unresolved. This paper is intended to show that if certain rather strong restrictions be placed on the size of the degrees of only the first two indeterminates then a polynomial n-cocycle is merely a coboundary. These restrictions are that the degree be less than p in the first indeterminate and less than $p-1$ in the second.

There is a close connection between polynomial cocycles and the n'th cohomology group $H^n(\pi, G)$ where π is a finite abelian group of exponent p (see [6], pp. 104 and 105) and G is any subgroup of π upon which π is taken to operate simply. Since π is isomorphic to the additive group $GF(p^r)^+$ of a Galois field of order p^r for some r and similarly for G, we can consider that π and G are additive groups of Galois fields, one being an algebraic extension of the other. It is easy to show that the functions on n variables of a Galois field into itself are polynomials over the field. Therefore, knowing the structure of the polynomial cocycles, gives us explicitly representatives of the cosets of $Z^n(\pi, G)$ with respect to $B^n(\pi, G)$, these being respectively the group of n-cocycles and n-coboundaries of π over G.

2. PRELIMINARIES. If R is a ring with an identity, let D be a endomorphism of the additive group of the ring of polynomials in n-indeterminates over R given by

$$Df(x_1, \ldots, x_n) = f(x_1 + 1, x_2, \ldots, x_n) - f(x_1, \ldots, x_n). \tag{1}$$

Let $\delta f(x_1, \ldots, x_{n+1})$ denote the coboundary of the polynomial $f(x_1, \ldots, x_n)$. Then by a straight forward computation we obtain:

PROPOSITION 1.

$$D(\delta f)(x_1, \ldots, x_{n+1}) = \delta(Df)(x_1, \ldots, x_{n+1}) - Df(x_2, \ldots, x_{n+1}). \tag{2}$$

We shall say that a polynomial is *normalized* in a certain indeterminate if the polynomial is zero whenever zero is substituted for that indeterminate.

PROPOSITION 2. *If K is a field of characteristic p and m an integer with $0 \leqq m \leqq p-2$, then there is a normalized polynomial $\psi_m(x)$ over K such that $D\psi_m(x) = x^m$.*

PROOF. Let $\psi_0(x) = x$. Clearly $\psi_0(x)$ satisfies the conclusion for $m=0$. Suppose the proposition true for all integers i of K with $0 < i < m \leqq p-2$. Now

$$D x^{m+1} = (m+1)\, x^m + \sum_{\nu=0}^{m-1} \binom{m+1}{\nu} x^\nu .$$

By the inductive assumption, there are normalized polynomials $\psi_\nu(x)$ over K with $D\psi_\nu(x) = x^\nu$, $0 \leqq \nu \leqq m-1$. Then

$$x^m = D\left[\frac{1}{m+1}\left(x^{m+1} - \sum_{\nu=0}^{m-1}\binom{m+1}{\nu}\psi_\nu(x)\right)\right],$$

and we let $\psi_m(x)$ be the polynomial in brackets.

At a suggestion of Professor WHAPLES, a formula due to LYNDON, see [5], p. 732, formula 5a, was useful in developing the following:

PROPOSITION 3. *Let K be a field of characteristic p and f a polynomial $(n+1)$-cocycle over K, normalized in the first two indeterminates, of degree less than p in the first and less than $p-1$ in the second. Further let g be a polynomial in n indeterminates over K, normalized and of degree less than p in the first indeterminate and such that*

$$Dg(x_1, \ldots, x_n) = -f(1, x_1, \ldots, x_n). \tag{3}$$

Then f is the coboundary of g.

PROOF. From (3) it follows that

$$g(1, x_2, \ldots, x_n) - g(0, x_2, \ldots, x_n) = -f(1, 0, x_2, \ldots, x_n).$$

Thus by the normalization assumptions on f and g, we find that $g(1, x_2, \ldots, x_n) = 0$. If c is any element of K, let $\delta g(c, x_1, \ldots, x_n)$ be the coboundary taken with c treated as variable in the expansion. Then

$$\begin{aligned}\delta g(1, x_1, \ldots, x_n) &= g(x_1, \ldots, x_n) - g(1 + x_1, x_2, \ldots, x_n) + \\ &\quad + \sum_{\nu=1}^{n-1} (-1)^{\nu+1} g(1, x_1, \ldots, x_{\nu-1}, x_\nu + x_{\nu+1}, x_{\nu+2}, \ldots, x_n) + \\ &\quad + (-1)^{n+1} g(1, x_1, \ldots, x_{n-1}) \\ &= -Dg(x_1, \ldots, x_n) = f(1, x_1, \ldots, x_n).\end{aligned}$$

Suppose for some integer m with $1 < m \leqq p-2$

$$\delta g(m, x_1, \ldots, x_n) = f(m, x_1, \ldots, x_n). \tag{4}$$

By Proposition 1

$$\delta g(m+1, x_1, \ldots, x_n) = \delta g(m, x_1, \ldots, x_n) + \delta(Dg)(m, x_1, \ldots, x_n) - Dg(x_1, \ldots, x_n).$$

Using (3) and the inductive assumption, yields

$$\delta g(m+1, x_1, \ldots, x_n) = f(m, x_1, \ldots, x_n) + \delta(Dg)(m, x_1, \ldots, x_n) + f(1, x_1, \ldots, x_n).$$

But

$$\begin{aligned}\delta(Dg)(m, x_1, \ldots, x_n) &= Dg(x_1, \ldots, x_n) - Dg(m + x_1, x_2, \ldots, x_n) + \\ &\quad + \sum_{\nu=1}^{n-1} (-1)^{\nu+1} Dg(m, x_1, \ldots, x_{\nu-1}, x_\nu + \\ &\quad + x_{\nu+1}, x_{\nu+2}, \ldots, x_n) + (-1)^{n+1} Dg(m, x_1, \ldots, x_n) \\ &= -f(1, x_1, \ldots, x_n) + \delta f(1, m, x_1, \ldots, x_n) + \\ &\quad + f(m+1, x_1, \ldots, x_n) - f(m, x_1, \ldots, x_n).\end{aligned}$$

f is an $(n+1)$-cocycle, so that $\delta f(1, m, x_1, \ldots, x_n) = 0$, and (4) holds with m replaced by $m+1$. Hence by induction (4) holds for all non-zero integers of K. Since f and g are normalized in the first indeterminate, (4) also holds when $m = 0$. If we let

$$\begin{aligned}h(x_1, \ldots, x_{n+1}) &= f(x_1, \ldots, x_{n+1}) - \delta g(x_1, \ldots, x_{n+1}) \\ &= \sum_{\nu=1}^{s} x_1^\nu h_\nu(x_2, \ldots, x_{n+1}).\end{aligned}$$

where s is the degree in x_1, then $s \leqq p-1$ and $h(m, x_2, \ldots, x_{n+1}) = 0$ for $m = 0, 1, \ldots, p-1$. Clearly all h_ν are identically zero, and the proposition follows.

The foregoing proposition can be generalized somewhat in that it holds for polynomials over a commutative integral domain with an identity, p being taken as the additive order of the identity which may of course be infinite.

3. Results.

Theorem. *For $n \geqq 2$ and K a field of characteristic p, every polynomial n-cocycle over K of degree less than p in the first indeterminate and less than $p-1$ in the second is a coboundary.*

Proof. First suppose f is a polynomial n-cocycle normalized in the first two indeterminates. Then

$$f(1, x_1, \ldots, x_{n-1}) = \sum_{\nu=1}^{s} x_1^\nu g_\nu(x_2, \ldots, x_{n-1})$$

with $s \leqq p-2$. Let

$$g(x_1, \ldots, x_{n-1}) = -\sum_{\nu=1}^{s} \psi_\nu(x_1)\, g_\nu(x_2, \ldots, x_{n-1})$$

where the $\psi_\nu(x_1)$ are those normalized polynomials of proposition 2 with $D\psi_\nu(x_1) = x_1^\nu$. Clearly

$$Dg(x_1, \ldots, x_{n-1}) = -f(1, x_1, \ldots, x_{n-1}),$$

and f is the coboundary of g by proposition 3.

If f is as above but not normalized in the first two indeterminates, then by the method used in proving lemma 6.1 of [1], observing that the functions used there are now polynomials, it can be shown that f is cohomologous to a polynomial n-cocycle, normalized in these first two indeterminates and of degree no larger than that of f in these indeterminates. This statement can also easily be proved directly without recourse to lemma 6.1 of [1]. The theorem clearly follows.

REMARKS. As was done in [2], the method used in the proof of the above theorem can also be used to show that over a field of characteristic zero all polynomial cocycles are coboundaries; for in characteristic zero, propositions 2 and 3 go through with no restriction on the degree. This, however, was proved by a much shorter method in [4].

In characteristic p it is not possible to relax the requirements given in the above theorem on the degree of the first two indeterminates and always expect a polynomial cocycle with the larger degrees in these indeterminates to be a coboundary. To see this, let $\varphi_p(x, y)$ be the image of the polynomial $p^{-1}(x^p-(x+y)^p+y^p)$ under the natural homomorphism of $Z[x, y]$ onto $Z/(p)[x, y]$ where Z is the ring of rational integers. Then $\varphi_p(x, y)$ is a polynomial 2-cocycle over $GF(p)$ which is of degree $p-1$ in x and y but is not a coboundary. In fact, for any even integer $n\geqq 2$, the polynomial $c\varphi_p(x_1, x_2)\cdots\varphi_p(x_{n-1}, x_n)$ where c is in $GF(p)$ and all factors are of the form $\varphi_p(x_i, x_{i+1})$ is a n-cocycle which is not a coboundary. That it is a cocycle, follows from the well known coboundary formula for cup products, see (4.6) of [1], and a rather tedious argument on polynomial identities shows that it is not a coboundary. It is easy to see using (16.2) and (16.3) of [1] that these polynomials with $c=0, 1, \ldots, p-1$ form a complete set of representatives for the cosets of $Z^n(GF(p)^+, GF(p)^+)$ with respect to $B^n(GF(p)^+, GF(p)^+)$ where $GF(p)^+$ operates simply on itself. If $n>1$ is odd in the above, then the corresponding coset representatives are $cx_1\varphi_p(x_2, x_3)\cdots\varphi_p(x_{n-1}, x_n)$ where all factors after the first are $\varphi_p(x_i, x_{i+1})$. Again these polynomial cocycles are not coboundaries when $c\neq 0$. Also it can be shown that $x_1^{p^{r_1}}x_2^{p^{r_2}}x_3^{p^{r_3}}\cdots x_n^{p^{r_n}}$, $r_i\neq r_j$ and $r_i\neq 1$, is a polynomial n-cocycle of degree p in x_1 which is not a coboundary.

There are other examples similar to those just given. In fact, if we have a polynomial which is the product of the additive polynomials $x_i^{p^{r_i}}$ with $r_i\neq r_j$ when $i\neq j$ and the special polynomial 2-cocycles $\varphi_p(x_k, x_{k+1})^{p^{s_k}}$, then it is a cocycle but not a coboundary. For example, $x_1^{p^q}\varphi_p(x_2, x_3)^{p^r}x_4^{p^s}\varphi_p(x_5, x_6)^{p^t}$ where $q\neq s$ is a polynomial 6-cocycle which is not a coboundary. One could conjecture that if $n>2$ all polynomial n-cocycles which are not coboundaries are composed of a coboundary and a linear combination over the field of these special products. The main result of [3] shows that this statement is true at least when $n=3$. The requirement that $q\neq s$ in the above 6-cocycle is necessary in order that it not be a coboundary when $p\neq 2$. For suppose $q=s$ and $p\neq 2$. From proposition 2 of [3], $x_1^{p^s}\varphi_p(x_3, x_4)^{p^r}-\varphi_p(x_2, x_3)^{p^r}x_4^{p^s}$

is a coboundary, and it follows that the above example with $q=s$ is cohomologous to $x_1^{p^s} x_2^{p^s} \varphi_p(x_3, x_4)^{p^r} \varphi_p(x_5, x_6)^{p^t}$. But it is easy to check that this is the coboundary of $\frac{1}{2} x_1^{2p^s} \varphi_p(x_2, x_3)^{p^r} \varphi_p(x_4, x_5)^{p^t}$.

References

[1] Eilenberg, S., and S. MacLane: Cohomology Theory in Abstract Groups. I. Ann. Math. **48**, 51–78 (1947).

[2] Heaton, R.: Polynomial Cocycles. Doctoral Diss., Indiana University 1955.

[3] Heaton, R.: Polynomial 3-Cocycles over Fields of Characteristic p. Duke Math. J. **26**, 269–275 (1959).

[4] Heaton, R., and G. Whaples: Polynomial Cocycles. Duke Math. J. **25**, 691–696 (1958).

[5] Lyndon, R.: New Proof for a Theorem of Eilenberg and MacLane. Ann. Math. **50**, 731–735 (1949).

[6] Zassenhaus, H.: Lehrbuch der Gruppentheorie. Leipzig u. Berlin 1937.

Rutgers University, New Brunswick, New Jersey (U.S.A.)

(Eingegangen am 13. Oktober 1960)

ANDRÉ, J.
Math. Zeitschr. 76, 240—256 (1961)

Über Parallelstrukturen*)

Teil III: Zentrale t-Strukturen[1])

Von
JOHANNES ANDRÉ

Einleitung

Eine affine Parallelstruktur (a.P.) heißt *zentral*, wenn jede ihrer Translationen zentral ist, d.h. mit einer Geraden auch alle zu ihr parallelen festläßt. Das braucht, wie in II, §2 an Beispielen gezeigt wurde, bei Parallelstrukturen — im Gegensatz zu Ebenen — keineswegs immer der Fall zu sein (nicht einmal bei transitiver Translationsgruppe), muß vielmehr zusätzlich gefordert werden. Eine *zentrale t-Struktur* oder kurz *t-Struktur* ist eine zentrale a.P. mit „genügend vielen" Translationen, genauer wird die Existenz nichtidentischer Translationen in mindestens zwei verschiedenen Richtungen gefordert. Eine t-Ebene [s. ANDRÉ (1958, a)] ist eine t-Struktur, die zugleich affine Ebene ist.

Die Translationsgruppe einer t-Struktur ist abelsch. Jedes ihrer Transitivitätsgebiete (d.h. jedes Gebiet, das aus einem ihrer Punkte durch Anwendung aller Translationen entsteht) läßt sich in natürlicher Weise als eine zentrale Translationsstruktur (z.T.; vgl. II, §2) auffassen, indem man als Geraden die aus mindestens zwei Punkten bestehenden Schnitte der Geraden unserer t-Struktur mit dem fraglichen Transitivitätsgebiet nimmt (s. §2). Alle die so gewonnenen t-Strukturen sind paarweise isomorph (Satz 2.3). Man kann auf diese Weise jeder t-Struktur $\mathfrak{p}$ eindeutig eine z.T. mit derselben Translationsgruppe wie $\mathfrak{p}$ zuordnen, die man den *Stamm* von $\mathfrak{p}$ nennt.

Es liegt jetzt nahe, t-Strukturen mit vorgegebenem Stamm zu konstruieren; das läuft im wesentlichen auf die Konstruktion einer t-Struktur mit gegebener Translationsgruppe hinaus. Anders gewendet heißt das, eine vorgegebene z.T. so zu erweitern, daß die Translationsgruppe dabei erhalten bleibt (translationstreue Erweiterung, vgl. I, §4). Solche Erweiterungen sind auf mannigfache Weise möglich. Man kann sogar eine beliebige z.T. translationstreu in eine affine Ebene einbetten (die dann notwendig eine t-Ebene ist), wie dies in den §§ 3, 4 und 5 näher ausgeführt wird. Die fragliche Konstruktion geschieht durch eine Art „freie Erweiterung" der z.T., die an ein Verfahren von M. HALL (1943) zur Konstruktion freier Ebenen angelehnt ist. Die einzelnen Konstruktionsschritte sind hier aber komplizierter als bei HALL, denn es muß

*) Teil einer von der Naturwissenschaftlich-Philosophischen Fakultät der Technischen Hochschule Braunschweig angenommenen Habilitationsschrift.

[1]) Teil I, Math. Z. **76**, 85—102, Teil II, Math. Z. **76**, 155—163 (1961), im folgenden hier mit I bzw. II zitiert.

ja darauf geachtet werden, daß die Translationsgruppe bei den Erweiterungsschritten sich nicht ändert. In der zweiten Hälfte des §3 wird diese „freie translationstreue Erweiterung" kurz skizziert, in §4 ausführlich begründet, während wir in §5 schließlich zeigen, daß die so erhaltene „freie translationstreue Erweiterungsebene" tatsächlich alle gewünschten Eigenschaften hat (insbesondere auch die gegebene Gruppe zur Translationsgruppe). Wir haben damit, grob gesagt, gezeigt, daß es zu einer vorgegebenen abelschen Translationsgruppe stets affine Ebenen mit dieser Gruppe gibt[2]). Im letzten Paragraphen endlich werden noch die Streckungen dieser freien Erweiterungsebene untersucht. Es zeigt sich, daß auch sie mit denjenigen ihres Stammes übereinstimmen.

§ 1. Grundbegriffe

1. DEFINITION. Eine *zentrale t-Struktur* oder kurz *t-Struktur* ist eine a.P. $\mathfrak{p}=(P, \mathfrak{G})$ mit folgenden zusätzlichen Eigenschaften:

(z 1) *Jede Translation τ von $\mathfrak{p}$ ist zentral, d.h.*[3]) *aus $G^\tau = G$ und $G' \| G$ $(G, G' \in \mathfrak{G})$ folgt $G'^\tau = G'$.*

(z 2) *Es gibt in $\mathfrak{p}$ (mindestens) zwei nichtparallele Geraden G_1, G_2 und nichtidentische Translationen τ_1, τ_2 derart, daß $G_i^{\tau_i} = G_i$ $(i=1, 2)$ ist.*

BEMERKUNGEN. Aus (z 1) folgt nach I, Satz 4.6 (b), daß die Translationen einer t-Struktur eine Gruppe bilden, die wir wieder mit T bezeichnen wollen. Axiom (z 2) besagt, daß T in gewissem Sinne „genügend reichhaltig" ist. Genauer gesagt, bedeutet (z 2): *Es gibt auf der uneigentlichen Geraden von $\mathfrak{p}$ mindestens zwei Punkte, die Zentren nichtidentischer Translationen sind.* Mithin bedeutet (z 2), daß es nichtidentische Translationen in (mindestens) zwei verschiedenen „Richtungen" gibt. Die t-Ebenen [im Sinne von ANDRÉ (1958, a)] kann man kennzeichnen als affine Ebenen, die zugleich t-Strukturen sind.

Seien $\tau_1, \tau_2 \neq \underline{1}$ Translationen von $\mathfrak{p}$ mit den (uneigentlichen) Zentren z_1 bzw. $z_2 \neq z_1$. Dann hat $\tau_1 \tau_2 \neq \underline{1}$ weder z_1 noch z_2 zum Zentrum, denn die Translationen mit gleichem Zentrum bilden eine Untergruppe von T. Wegen (z 1) gilt also [in Verschärfung von (z 2)]:

SATZ 1.1. *Auf der uneigentlichen Geraden einer z.P. gibt es mindestens drei verschiedene Punkte, die Zentren nichtidentischer Translationen sind*[4]).

2. Spezielle t-Strukturen sind die sog. *T-Strukturen;* sie sind durch das folgende (z 2) verschärfende zusätzliche Axiom gekennzeichnet:

[2]) Eine kurze Zusammenfassung eines Teiles dieser Ergebnisse wurde bereits in ANDRÉ (1958, b) gebracht.

[3]) Diese Definition einer zentralen Translation steht im Einklang mit der in I, S. 97 gebrachten. Aus (z 1) folgt nämlich für jede Translation τ die Existenz einer Geraden G mit $G^\tau = G$. O.B.d.A. können wir $\tau \neq \underline{1}$ annehmen; dann ist $G = \langle p, p^\tau \rangle$ die gewünschte Gerade; das ergibt Bedingung (a) in I. Dadurch ist G im wesentlichen eindeutig bestimmt. Genauer: Ist G' eine weitere Gerade mit $G'^\tau = G'$ (wieder $\tau \neq \underline{1}$ angenommen), so $G \| G'$. Wäre nämlich $G' \# G$ und $q \in G'$, so wäre $q^\tau \notin G'$, also $G'^\tau \neq G'$. Damit ist auch Bedingung (b) bewiesen.

[4]) Die Anzahl „drei" kann allgemein nicht erhöht werden, wie das Beispiel der Ebene über dem Galoisfeld der Ordnung 2 zeigt.

(z 2') *Zu jeder Geraden G von $\mathfrak{p}$ gibt es (mindestens) eine nichtidentische Translation τ mit $G^\tau = G$.*

Mithin bedeutet (z 2'), daß es nichtidentische Translationen in jeder Richtung gibt, m.a.W., daß jeder Punkt der uneigentlichen Geraden von $\mathfrak{p}$ Zentrum einer nichtidentischen Translation ist. Spezielle T-Strukturen sind u.a. die zentralen Translationsstrukturen (vgl. II, §2). Eine genauere Untersuchung der allgemeinen T-Strukturen folgt in Teil IV.

3. Wir haben schon festgestellt, daß die Translationen einer t-Struktur eine Gruppe bilden. Wir zeigen jetzt genauer:

SATZ 1.2. *Die Gruppe der Translationen einer t-Struktur ist abelsch.*

BEWEIS. Zunächst seien τ und τ' nichtidentische Translationen mit verschiedenen Zentren; dann folgt $\tau\tau' = \tau'\tau$ aus I, Satz 4.5. Haben τ und τ' gleiche Zentren, so schließen wir die Vertauschbarkeit genau ebenso wie beim Beweis von II, Satz 2.3 mit Hilfe eines $\tau'' \neq \underline{1}$, das ein anderes Zentrum als τ hat. [Ein solches τ'' existiert nach (z 2).]

Aus der Definition von „translationstreu" (I, §4, Ziffer 5) folgt sofort:

SATZ 1.3. *Jede translationstreue Unterstruktur einer t- bzw. T-Struktur ist ebenfalls eine t- bzw. T-Struktur.*

Der folgende über Axiom (a 3) hinausgehende Reichhaltigkeitssatz wird uns später nützlich sein.

SATZ 1.4. *Zu jedem Punkt p einer t-Struktur $\mathfrak{p}$ kann man drei weitere Punkte von $\mathfrak{p}$ finden, die zusammen mit p ein nichtausgeartetes Viereck bilden (d.h. von den vier Punkten liegen keine drei auf einer Geraden).*

BEWEIS. Sei p ein beliebiger Punkt der t-Struktur $\mathfrak{p}$, ferner τ, τ' [nach (z 2) existierende] nichtidentische Translationen mit verschiedenen Zentren. Die vier Punkte p, p^τ, $p^{\tau'}$ und $p^{\tau\tau'}$ haben die gewünschte Eigenschaft, ein nichtausgeartetes Viereck zu bilden; denn p, p^τ und $p^{\tau'}$ liegen nicht auf einer Geraden, weil τ und τ' verschiedene Zentren haben. Da dann auch die Zentren von τ, $\tau\tau'$ bzw. τ', $\tau\tau'$ jeweils verschieden sind, liegen auch p, p^τ und $p^{\tau\tau'}$ bzw. p, $p^{\tau'}$, $p^{\tau\tau'}$ nicht auf einer Geraden. Schließlich gehören auch p^τ, $p^{\tau'}$ und $p^{\tau\tau'}$ nicht derselben Geraden an, denn sonst gelte dasselbe für die Punkte p, $p^{\tau'\tau^{-1}}$ und $p^{\tau'}$, die man aus den vorherigen durch Anwendung von τ^{-1} gewinnt. Aber auch τ' und $\tau'\tau^{-1}$ haben verschiedene Zentren.

ZUSATZ. *Durch jeden Punkt einer t-Struktur gehen mindestens drei Geraden.*

Schließlich gilt noch für T-Strukturen, über (a 4) hinausgehend:

SATZ 1.5. *Jede Gerade einer T-Struktur enthält mindestens zwei (eigentliche) Punkte.*

BEWEIS. Sei $\mathfrak{p} = (P, \mathfrak{G})$ eine T-Struktur und $G \in \mathfrak{G}$. Nach (a 4) gibt es einen Punkt $p \in G$, nach (z 2') eine Translation $\tau \neq \underline{1}$ mit $G^\tau = G$. Neben p gehört auch $p^\tau \neq p$ zu G.

§ 2. Der Stamm einer t-Struktur

1. Sei $\mathfrak{p}=(P, \mathfrak{G})$ eine t-Struktur und T ihre Translationsgruppe. Ist T transitiv, so ist $\mathfrak{p}$ eine zentrale Translationsstruktur; wir haben sie in II behandelt. Unabhängig von jeder Transitivitätsvoraussetzung können wir aber durch $\mathfrak{p}$ auf T in folgender Weise eine Partition $\mathfrak{P}$ definieren: Sei $\mathsf{T}(G)\subseteqq\mathsf{T}$ die Gruppe aller G (als Ganzes) festlassenden Translationen, dann soll $\mathfrak{P}$ aus genau denjenigen Untergruppen $\mathsf{T}(G)$ bestehen, die nicht nur die Identität enthalten[5]:

$$\mathfrak{P}=\{\mathsf{T}(G)\,|\,G\in\mathfrak{G},\ \mathsf{T}(G)>\{\underline{1}\}\}. \tag{2.1}$$

Da einerseits $\mathsf{T}(G)\cap\mathsf{T}(G')=\{\underline{1}\}$ für $\mathsf{T}(G)\neq\mathsf{T}(G')$ (also $G\nparallel G'$) ist und es andererseits zu jedem $\tau\in\mathsf{T}$ ein $G\in\mathfrak{G}$ mit $G^{\tau}=G$ gibt (s. Fußnote 3), sind die Bedingungen für eine Partition (vgl. II, §1) alle erfüllt. Aus T und $\mathfrak{P}$ gewinnen wir aber gemäß den Ergebnissen von II, §1 eine Translationsstruktur $\mathfrak{t}_{\mathfrak{p}}=[\mathsf{T}, \mathfrak{P}]$ (indem wir als Punkte die Elemente von T und als Geraden die Rechtsrestklassen nach den Untergruppen von $\mathfrak{P}$ nehmen), deren Translationsgruppe zu T isomorph ist (und die wir deshalb mit T identifizieren können und wollen); sie ist zentral, weil T abelsch ist (II, Satz 2.4). Die so gewonnene durch $\mathfrak{p}$ eindeutig bestimmte z.T. $\mathfrak{t}_{\mathfrak{p}}$ wird als *Stamm* von $\mathfrak{p}$ bezeichnet. Statt $\mathfrak{t}_{\mathfrak{p}}$ schreiben wir häufig einfach $\mathfrak{t}$, wenn kein Mißverständnis zu befürchten ist.

2. Wir können aus $\mathfrak{p}=(P, \mathfrak{G})$ noch auf eine andere vernünftige Weise z.T. gewinnen (von denen aber gezeigt wird, daß sie zum Stamm isomorph sind). Die folgende Bemerkung soll darauf vorbereiten. Durch die Translationsgruppe T zerfällt P in (endlich oder unendlich viele) Transitivitätsgebiete $T_i=p_i^{\mathsf{T}}$ ($p_i\in T_i$; $p_i\notin T_j$ für $i\neq j$; hierbei gehören i, j einer Indexmenge I an). Also ist

$$P=\bigcup_{i\in I} T_i \tag{2.2}$$

mit paarweise fremd gelegenen T_i.

Durch $\mathfrak{p}$ werden auf den T_i in natürlicher Weise geometrische Strukturen $\mathfrak{t}_i$ aufgeprägt: Sei $G\in\mathfrak{G}$; der Durchschnitt $G\cap T_i$ kann leer sein; ist er es nicht, so nennen wir ihn die *Spur* der Geraden G auf T_i. Die geometrische Struktur $\mathfrak{t}_i=(T_i, \mathfrak{L}_i)$ [die von $\mathfrak{p}$ auf T_i *induzierte Struktur*] definieren wir jetzt auf folgende Weise: Ihre *Punkte* sind die Elemente von T_i; ihre *Geraden* (d.h. die Elemente aus $\mathfrak{L}_i$) die aus mindestens zwei Elementen bestehenden Spuren der Geraden aus $\mathfrak{G}$ auf T_i, d.h.

$$\mathfrak{L}_i=\{G\cap T_i\,|\,G\in\mathfrak{G},\ |G\cap T_i|>1\}; \tag{2.3}$$

zwei Geraden $G\cap T_i$ und $G'\cap T_i$ nennen wir in $\mathfrak{t}_i$ zueinander *parallel*, wenn G und G' es in $\mathfrak{p}$ sind.

Wenn wir zeigen, daß jede dieser induzierten Strukturen $\mathfrak{t}_i$ zum Stamm von $\mathfrak{p}$ isomorph ist (d.h. eine eineindeutige, inzidenz- und parallelenerhaltende Abbildung zwischen den Punkten und Geraden von $\mathfrak{t}_i$ und denen von $\mathfrak{t}_{\mathfrak{p}}=\mathfrak{t}$

[5]) Offensichtlich gilt $\mathsf{T}(G)=\mathsf{T}(G')$ für $\mathsf{T}(G)>\{\underline{1}\}$ genau dann, wenn $G\|G'$ ist, denn $\mathfrak{p}$ ist als zentral vorausgesetzt (vgl. auch Fußnote 3).

besteht), so wissen wir damit auch, daß ι_i eine z.T. ist, denn der Stamm ι von $\mathfrak{p}$ war als solche erkannt.

Bevor wir dies beweisen, müssen wir noch zwei andere, auch für sich interessante und wichtige Sätze herleiten.

SATZ 2.1. *Ist $G \cap T_i \neq \emptyset$ (d.h. existiert die Spur von G auf T_i), so stimmen die Mächtigkeiten von $G \cap T_i$ und der Gruppe $\mathsf{T}(G)$ aller G festlassenden Translationen überein:* $|G \cap T_i| = |\mathsf{T}(G)|$.

BEWEIS. Sei p fest auf $G \cap T_i$ gewählt. Aus $q \in G \cap T_i$ folgt die eindeutige Existenz eines $\tau \in \mathsf{T}$ mit $p^\tau = q$, da p und q in demselben Transitivitätsgebiet T_i liegen. Dieses τ läßt G fest, d.h. $\tau \in \mathsf{T}(G)$. Umgekehrt folgt hieraus $p^\tau \in G \cap T_i$. Durch $\tau \to p^\tau$ ist eine eineindeutige Zuordnung zwischen den Elementen von $\mathsf{T}(G)$ und $G \cap T_i$ hergestellt.

BEMERKUNG. Aus diesem Satz folgt insbesondere, daß die Mächtigkeit von $G \cap T_i$ nicht von der speziellen Wahl von i abhängt, wenn nur $G \cap T_i$ nicht leer ist.

Ist $\mathsf{T}(G) = \{\underline{1}\}$ für ein $G \in \mathfrak{G}$, d.h. gibt es keine G festlassende nichtidentische Translation (also keine in der „Richtung" G), so ist $G \cap T_i$ leer oder ein Punkt, der *Spurpunkt* von G auf T_i. Für $\mathsf{T}(G) > \{\underline{1}\}$ ist dagegen $G \cap T_i$ leer oder besteht aus mindestens zwei Punkten. Hieraus folgt sofort:

SATZ 2.2. *Eine t-Struktur $\mathfrak{p} = (P, \mathfrak{G})$ ist genau dann T-Struktur, wenn für alle $G \in \mathfrak{G}$ und alle $i \in I$ entweder $G \cap T_i = \emptyset$ oder $|G \cap T_i| > 1$ gilt, d.h. wenn alle Spuren mindestens zwei Punkte enthalten.*

3. Nunmehr wollen wir eine isomorphe Abbildung vom Stamm ι von $\mathfrak{p}$ auf eine induzierte Struktur ι_i herstellen. Die Punkte von ι sind die Translationen von $\mathfrak{p}$. Der identischen Translation $\underline{1}$ möge ein beliebiger, aber dann festgehaltener Punkt p_i von ι_i zugeordnet sein. Dann vermittelt die eineindeutige Zuordnung

$$\iota_i: \quad \tau \to p_i^\tau \qquad (\tau \in \mathsf{T}) \tag{2.4}$$

den gewünschten Isomorphismus zwischen ι und ι_i. Die Geraden von ι sind nämlich von der Gestalt $\mathsf{T}(G)\tau$ mit $\tau \in \mathsf{T}$, $G \in \mathfrak{G}$ und $\mathsf{T}(G) > \{\underline{1}\}$, wobei wir o.B.d.A. $p_i \in G$ annehmen können. [Sonst ersetze man G durch $\langle G \| p_i \rangle$, was wegen $\mathsf{T}(G) = \mathsf{T}(\langle G \| p_i \rangle)$ zulässig ist[6].] Wegen Satz 2.1 und der Geradendefinition in ι ist $|G \cap T_i| = |\mathsf{T}(G)| > 1$, mithin $G \cap T_i$ eine Gerade von ι_i. Durch ι_i wird die Gerade $\mathsf{T}(G)$ von ι auf die Gerade $G \cap T_i$ von ι_i abgebildet, also $\mathsf{T}(G)\tau$ auf $(G \cap T_i)^\tau = G^\tau \cap T_i$. Die Parallelität bleibt bei Anwendung von ι_i erhalten: Zwei Geraden $\mathsf{T}(G)\tau$ und $\mathsf{T}(G')\tau'$ ($\tau, \tau' \in \mathsf{T}$; $G, G' \in \mathfrak{G}$; $p_i \in G \cap G'$) aus ι sind nämlich genau dann zueinander parallel, wenn $G \| G'$, mithin $G = G'$ ist. Dann folgt aber $G^\tau \| G^{\tau'}$ und somit (wegen der Definition der Parallelität in ι_i) auch $G^\tau \cap T_i \| G^{\tau'} \cap T_i$. Der durch (2.4) definierte (und im wesentlichen eindeutig bestimmte[7]) Isomorphismus ι_i von ι auf ι_i ($i \in I$) wird als ein *kanonischer Isomorphismus* bezeichnet.

[6] Als $\langle G \| p \rangle$ wurde in I die durch p gehende und zu G parallele Gerade definiert.

[7] Durch die Vorgabe von p_i, des Bildpunktes von $\underline{1}$, ist ι_i eindeutig bestimmt. Wird statt dessen $\underline{1}$ in $q_i = p_i^\tau$ ($\tau \in \mathsf{T}$) abgebildet, so erhalten wir $\iota_i \tau$ als zugehörigen kanonischen Isomorphismus. Bis auf Translationen ist er also eindeutig bestimmt.

Damit ist die Isomorphie zwischen t und t_i hergestellt, insbesondere auch gezeigt, daß die t_i z.T. sind. Man kann jedes t_i als Unterstruktur (im Sinne von I, §2, Ziffer 5) von $\mathfrak{p}$ auffassen, indem man $G \cap T_i$ mit G identifiziert. [Diese Identifizierung ist möglich, denn aus $G \cap T_i = G' \cap T_i$ $(G, G' \in \mathfrak{G})$ mit $|G \cap T_i| > 1$ folgt $G = G'$.] Da offensichtlich [im Sinne von (a') und (b') von I, §4, Ziffer 5] die Translationsgruppen von $\mathfrak{p}$ und t_i übereinstimmen (nämlich $= \mathsf{T}$ sind), ist t_i sogar translationstreue Unterstruktur von $\mathfrak{p}$. Da wir ferner (wegen der bewiesenen Isomorphie) jedes t_i mit dem Stamm $t_\mathfrak{p}$ identifizieren können, läßt sich jede t-Struktur auch als translationstreue Erweiterung ihres Stammes auffassen. Zusammen erhalten wir:

Satz 2.3. *Jedes Transitivitätsgebiet T_i $(i \in I)$ bezüglich der (vollen) Translationsgruppe T einer t-Struktur $\mathfrak{p}$ ist eine zentrale Translationsstruktur t_i hinsichtlich T, wenn man als Geraden von t_i die aus mindestens zwei Punkten bestehenden Spuren der Geraden von $\mathfrak{p}$ nimmt und zwei Geraden von t_i genau dann parallel nennt, wenn ihre erzeugenden Geraden aus $\mathfrak{p}$ diese Eigenschaft haben.*

Jedes dieser t_i ist zum Stamm $t_\mathfrak{p}$ von $\mathfrak{p}$ isomorph, insbesondere sind also die t_i auch untereinander isomorph. Schließlich ist $\mathfrak{p}$ translationstreue Erweiterung eines jeden t_i und auch des Stammes $t_\mathfrak{p}$.

Ist $\mathfrak{p}$ endlich, so auch die t_i. Aus der Isomorphie der t_i untereinander folgt insbesondere $|T_i| \mid |P|$:

Satz 2.4. *Die Anzahl der Punkte des Stammes $t_\mathfrak{p}$ einer endlichen t-Struktur $\mathfrak{p}$ ist Teiler der Anzahl der Punkte von $\mathfrak{p}$.*

4. Wir geben jetzt noch eine Übersicht über translationstreue Unterstrukturen, die Translationsstrukturen sind.

Satz 2.5. *Die translationstreuen Unterstrukturen einer t-Struktur $\mathfrak{p}$, die Translationsstrukturen sind, stimmen genau mit den von $\mathfrak{p}$ induzierten Strukturen t_i überein.*

Beweis. Jedenfalls sind die t_i nach Satz 2.3 translationstreue Unterstrukturen. Sei umgekehrt $t = (T, \mathfrak{L})$ eine in $\mathfrak{p} = (P, \mathfrak{G})$ translationstreu gelegene Translationsstruktur (die nach Satz 1.2 und II, Satz 2.4 zentral ist), insbesondere gilt $T \subseteq P$ und $\mathfrak{L} \subseteq \mathfrak{G}$. Da t translationstreu in $\mathfrak{p}$ liegt, ist $T^\tau = T$ für jede Translation τ von $\mathfrak{p}$, und umgekehrt kann jede Translation von t auf $\mathfrak{p}$ fortgesetzt werden. Somit ist T Transitivitätsgebiet hinsichtlich der Translationsgruppe T von $\mathfrak{p}$ (die mit der Translationsgruppe von t übereinstimmt); denn t war als Translationsstruktur vorausgesetzt. Es gibt also ein $i \in I$ mit $T = T_i$. Wir wollen jetzt $t = t_i$ $[= (T_i, \mathfrak{L}_i)]$ beweisen. Jedes $L \in \mathfrak{L}$ hat mit T_i mindestens zwei Punkte gemeinsam [das folgt aus Axiom (t 2)]. Die Gerade L, als Element von $\mathfrak{G}$ aufgefaßt, hat also auf T_i eine aus mindestens zwei Punkten bestehende Spur, also gilt $L \in \mathfrak{L}_i$, d.h. $\mathfrak{L} \subseteq \mathfrak{L}_i$. Ist umgekehrt $L \in \mathfrak{L}_i$, so hat L mit T_i mindestens zwei Punkte t, t' gemeinsam. Die Gerade $\langle t, t' \rangle_t \in \mathfrak{L}$ stimmt mit L überein, also gilt auch $\mathfrak{L}_i \subseteq \mathfrak{L}$, mithin $\mathfrak{L}_i = \mathfrak{L}$. Inzidenz und Parallelität in t zieht die in t_i nach sich und umgekehrt, also gilt $t = t_i$, w.z.b.w.

§ 3. Zentrale t-Strukturen mit vorgegebenem Stamm

Im vorigen Paragraphen haben wir gefunden, daß wir jeder t-Struktur $\mathfrak{p}$ eindeutig eine zentrale Translationsstruktur (z.T.) $\mathfrak{t}$, den Stamm von $\mathfrak{p}$ zuordnen können. Wir beginnen jetzt mit der Untersuchung des Umkehrproblemes: *t-Strukturen mit vorgegebenem Stamm zu konstruieren.* Sei $\mathfrak{p}$ eine t-Struktur mit dem Stamm $\mathfrak{t}$. Wir können, wie im vorigen Paragraphen ausgeführt wurde, $\mathfrak{p}$ als translationstreue Erweiterung von $\mathfrak{t}$ auffassen. Umgekehrt: Ist dies der Fall, so stimmt $\mathfrak{t}$ nach Satz 2.5 mit einer durch $\mathfrak{p}$ induzierten Translationsstruktur überein, ist also zum Stamm von $\mathfrak{p}$ isomorph. Mithin ist das Problem „Alle t-Strukturen mit vorgegebenem Stamm $\mathfrak{t}$ zu finden" äquivalent mit der Aufgabe „Alle translationstreuen Erweiterungen einer z.T. $\mathfrak{t}$ zu suchen, die t-Strukturen sind". Diese letztere Fragestellung kann man auch dahingehend verallgemeinern, daß die Voraussetzung, $\mathfrak{t}$ sei eine z.T., fallengelassen wird.

In Teil IV bringen wir eine vollständige Übersicht über alle möglichen translationstreuen Erweiterungen, die T-Strukturen sind. *Beispiele* von t-Strukturen, die translationstreu über $\mathfrak{t}$ sind, wurden bereits in I, §4, Ziffer 5 gebracht: In der Tat sind ja die dort eingeführten Strukturen $\mathfrak{t}^{(J)}$ translationstreu über $\mathfrak{t}$, und ist $\mathfrak{t}=(T, \mathfrak{L})$ eine z.T., so ist offensichtlich $\mathfrak{t}^{(J)}$ eine t-Struktur [dagegen sicher keine T-Struktur, denn keine der Geraden (t, j) $(t \in T, j \in J)$ bleibt bei einer nichtidentischen Translation fest.] Allerdings ist bei $\mathfrak{t}^{(J)}$ nur die Geradenmenge echt erweitert gegenüber der von $\mathfrak{t}$, während die Punktmengen beider Strukturen übereinstimmen.

Wir wollen in den folgenden Paragraphen zeigen, daß man zu einer beliebig vorgegebenen z.T. $\mathfrak{t}$ auch translationstreue Erweiterungen finden kann, deren Punktmengen *echte* Erweiterungen derjenigen von $\mathfrak{t}$ sind. [Dann gilt wegen (a 1) dasselbe auch für die Geradenmengen.] Weitergehend können wir sogar zeigen: *Jede zentrale Translationsstruktur kann translationstreu zu einer affinen Ebene erweitert werden.* Allgemeiner werden wir folgendes Problem lösen: Eine t-Struktur $\mathfrak{p}$ translationstreu in eine affine Ebene einzubetten. Eine solche Ebene können wir aus $\mathfrak{p}$ schrittweise durch eine „freie translationstreue Erweiterung[8])" von $\mathfrak{p}$ gewinnen, wie in den folgenden Paragraphen näher ausgeführt wird. Wir können hierzu natürlich annehmen, daß $\mathfrak{p}$ keine affine Ebene ist, also windschiefe Geraden (vgl. I, §2, Ziffer 2) enthält. Wir ordnen zunächst je zwei solchen Geraden einen neuen Schnittpunkt zu und erweitern $\mathfrak{p}$ auf diese Weise zu einer Inzidenzstruktur $\hat{\mathfrak{p}}$ (die keine Parallelstruktur ist). Die Geradenmenge von $\hat{\mathfrak{p}}$ stimmt mit derjenigen von $\mathfrak{p}$ überein, die Punktmenge von $\hat{\mathfrak{p}}$ ist aber eine echte Erweiterung. Die Translationen von $\mathfrak{p}$ lassen sich auf $\hat{\mathfrak{p}}$ in naturgemäßer Weise fortsetzen. Bei der nächsten Erweiterung von $\hat{\mathfrak{p}}$ auf eine Struktur $\mathfrak{p}_1$ bleibt die Punktmenge ungeändert; die Geradenmenge vergrößern wir auf folgende Weise: (1) Anwendung aller Translationen von $\mathfrak{p}$ mit festem (uneigentlichem) Zentrum z auf einen Punkt p von $\hat{\mathfrak{p}}$ ergibt

[8]) Sie ist analog zu der in HALL (1943) eingeführten „freien Erweiterung" von Inzidenzstrukturen, aber komplizierter, da hier bei den Erweiterungsschritten auch die Translationstreue erhalten bleiben muß.

eine Punktmenge, die wir als neue Gerade von $\mathfrak{p}_1$ nehmen, wenn sie nicht nur aus p allein besteht (was bedeuten würde, daß nur die identische Translation z zum Zentrum hätte), und wenn die Punktmenge nicht schon Teil einer Geraden aus $\hat{\mathfrak{p}}$ ist. (2) Jedes Paar verschiedener Punkte von $\hat{\mathfrak{p}}$, durch das noch keine bisher konstruierte Gerade geht, wird zu einer Geraden gemacht, die nur mit diesen beiden Punkten inzidiert. (3) Durch jeden Punkt wird zu jeder Geraden eine „Parallele" hinzugefügt, falls noch keine existiert. Diese neu gewonnene Gerade inzidiert nur mit diesem Punkt.

· Die so gewonnene Erweiterung $\mathfrak{p}_1$ von $\mathfrak{p}_0=\mathfrak{p}$ ist translationstreu und, bei geeigneter Definition der Parallelität, wieder eine t-Struktur. Fortführung desselben Konstruktionsverfahrens ergibt eine Kette translationstreuer Erweiterungen $\mathfrak{p}=\mathfrak{p}_0\leqq\mathfrak{p}_1\leqq\mathfrak{p}_2\leqq\cdots$. Die Vereinigung $[\mathfrak{p}]=\bigcup_{i=0}^{\infty}\mathfrak{p}_i$ dieser Kette (vgl. I, §2, Ziffer 5) ist eine affine Ebene, die wir *freie translationstreue Erweiterungsebene* von $\mathfrak{p}$ nennen. Im nächsten Paragraphen werden wir den ersten Erweiterungsschritt, nämlich den von $\mathfrak{p}=\mathfrak{p}_0$ nach $\mathfrak{p}_1$, im einzelnen durchführen; im übernächsten werden wir durch Iteration dieses Verfahrens die Ebene $[\mathfrak{p}]$ bilden und zeigen, daß sie eine translationstreue Erweiterung von $\mathfrak{p}$ ist (womit dann auch bestätigt wird, daß die Bezeichnung „freie *translationstreue* Erweiterungsebene" zu Recht besteht).

§4. Der erste Erweiterungsschritt

1. Es sei $\mathfrak{p}_0=(P_0, \mathfrak{G}_0)$ eine t-Struktur und T deren Translationsgruppe. Über $\mathfrak{p}_0$ machen wir noch die beiden folgenden Voraussetzungen:

(A) *$\mathfrak{p}_0$ ist keine affine Ebene.*

(B) *In $\mathfrak{p}_0$ gibt es mindestens fünf Geraden G_j ($j=1, \ldots, 5$), von denen keine zwei zueinander parallel sind, und für die $\mathsf{T}(G_j)=\{\underline{1}\}$ gilt.* (Das heißt auf der uneigentlichen Geraden von $\mathfrak{p}_0$ gibt es fünf Punkte, die Zentren nur von der identischen Translation sind.)

Für das zu untersuchende Problem, $\mathfrak{p}_0$ translationstreu in eine affine Ebene einzubetten, bedeuten diese Voraussetzungen keine Beschränkung der Allgemeinheit. Ist nämlich $\mathfrak{p}_0$ eine affine Ebene, so ist das Problem bereits trivialerweise gelöst. Ist (B) für $\mathfrak{p}_0$ nicht erfüllt, so betten wir $\mathfrak{p}_0$ translationstreu in ein $\mathfrak{p}_0^{(J)}$ mit $|J|\geqq 5$ ein (vgl. §3) und führen dann die Konstruktionsschritte für $\mathfrak{p}_0^{(J)}$ anstatt für $\mathfrak{p}_0$ durch. Die Voraussetzung (B) ist in $\mathfrak{p}_0^{(J)}$ z.B. für die Geraden $G_j=(p, j)$ ($p\in P_0$; $j=1, \ldots, 5$) erfüllt.

Wir definieren die Erweiterung $\hat{\mathfrak{p}}_0=(\hat{P}_0, \hat{\mathfrak{G}}_0)$ von $\mathfrak{p}_0$ auf folgende Weise durch Hinzunahme neuer Punkte, aber Beibehaltung der alten Geradenmenge: (α) $\hat{\mathfrak{G}}_0=\mathfrak{G}_0$; ($\beta$) $\hat{P}_0$ bestehe aus den Punkten von P_0 und der Gesamtheit aller Mengen $\{G, G'\}$ von in $\mathfrak{p}_0$ windschief gelegenen Geraden $G, G'\in\mathfrak{G}_0$:

$$\hat{P}_0=P_0\cup\{\{G, G'\}\,|\,G, G'\in\mathfrak{G}_0,\ G\cap G'=\emptyset,\ G\nparallel G'\}. \tag{4.1}$$

Der Punkt $\{G, G'\}$ soll mit G und G', jedoch mit keiner weiteren Geraden aus $\mathfrak{G}_0$ inzidieren, während die Punkte aus P_0 in $\hat{\mathfrak{p}}_0$ mit genau denjenigen Geraden

inzidieren, mit denen sie es in $\mathfrak{p}_0$ tun. Da $\mathfrak{p}_0$ keine affine Ebene ist, es in $\mathfrak{p}_0$ also Paare windschiefer Geraden gibt, ist $\hat{P}_0 > P_0$, also $\hat{\mathfrak{p}}_0$ eine echte Erweiterung von $\mathfrak{p}_0$. Offensichtlich ist $\hat{\mathfrak{p}}_0$ eine Inzidenzstruktur, denn durch zwei verschiedene Punkte von $\hat{P}_0$ geht höchstens eine Gerade aus $\hat{\mathfrak{p}}_0$. Dagegen ist $\hat{\mathfrak{p}}_0$ keine Parallelstruktur, denn (a 1) ist nicht erfüllt; vielmehr gilt:

HILFSSATZ 4.1. *Es gibt in* $\hat{P}_0 - P_0$ *mindestens zwei Punkte, durch die keine Verbindungsgerade aus* $\hat{\mathfrak{G}}_0 = \mathfrak{G}_0$ *geht.*

Zum Beweis brauchen wir wegen der in (β) gebrachten Inzidenzbedingung nur zu zeigen, daß es zwei Paare $\{G_1, G_1'\}$ und $\{G_2, G_2'\}$ von in $\mathfrak{p}_0$ windschief gelegenen Geraden gibt derart, daß G_1, G_1', G_2, G_2' paarweise verschieden sind. Zunächst gibt es *ein* Paar $\{G_1, G_1'\}$ windschiefer Geraden; ferner existiert nach Satz 1.1 eine Translation τ, deren Zentrum weder der uneigentliche Punkt von G_1 noch der von G_1' ist, also gilt $G_1^\tau \neq G_1$ und $G_1'^\tau \neq G_1'$, aber auch $G_1^\tau \neq G_1'$ und $G_1'^\tau \neq G_1'$ wegen $G_1 \nparallel G_1'$. Somit sind $\{G_1, G_1'\}$ und $\{G_1^\tau, G_1'^\tau\}$ die gewünschten Paare.

Jede (affine) Kollineation γ von $\mathfrak{p}_0$ läßt sich in naturgemäßer Weise eindeutig zu einer Kollineation von $\hat{\mathfrak{p}}_0$ fortsetzen (die wir ebenfalls γ nennen können und wollen), indem wir $\{G, G'\}^\gamma = \{G^\gamma, G'^\gamma\}$ setzen. (Dies können wir tun, denn liegen G und G' in $\mathfrak{p}_0$ zueinander windschief, so auch G^γ und G'^γ.) Diese Bemerkung gilt insbesondere auch für die Translationen von $\mathfrak{p}_0$.

2. Wir erweitern jetzt $\hat{\mathfrak{p}}_0 = (\hat{P}_0, \hat{\mathfrak{G}}_0)$ zu einer Struktur $\mathfrak{p}_1 = (P_1, \mathfrak{G}_1)$, indem wir nur die Geradenmenge $\hat{\mathfrak{G}}_0$ zu $\mathfrak{G}_1$ vergrößern, die Punktmenge aber ungeändert lassen: $P_1 = \hat{P}_0$. Dabei soll die Parallelitätsrelation von $\mathfrak{p}_0$ auf $\mathfrak{p}_1$ erweitert werden, und zwar so, daß die Parallelität auch in $\mathfrak{p}_1$ eine reflexive, symmetrische und transitive Relation ist. Die Menge $\mathfrak{G}_1$ der Geraden von $\mathfrak{p}_1$ soll aus den folgenden Elementen bestehen:

(0) Aus den Geraden von $\mathfrak{G}_0$. Zwei solche Geraden sollen in $\mathfrak{p}_1$ genau dann zueinander parallel liegen, wenn dies in $\mathfrak{p}_0$ der Fall ist.

(1) Aus den Mengen $p^{\mathsf{T}(G)}$ mit $p \in \hat{P}_0$ und $G \in \mathfrak{G}_0$, wenn (a) $\mathsf{T}(G) > \{1\}$ (d.h. die Menge $p^{\mathsf{T}(G)}$ mindestens zwei Punkte enthält) und (b) die Menge $p^{\mathsf{T}(G)}$ keine Teilmenge von Punkten einer Geraden aus $\mathfrak{G}_0$ ist. Inzidenz mit $p^{\mathsf{T}(G)}$ soll (mengentheoretische) Zugehörigkeit zu $p^{\mathsf{T}(G)}$ bedeuten. Die Menge der so erhaltenen Geraden $p^{\mathsf{T}(G)}$ bezeichnen wir mit $\mathfrak{G}_{01}$. Definitionsgemäß sei genau dann ein $G \in \mathfrak{G}_0$ zu einem $G = p^{\mathsf{T}(G')} \in \mathfrak{G}_{01}$ ($G' \in \mathfrak{G}_0$) parallel, wenn $G \| G'$ in $\mathfrak{p}_0$ ist, ferner seien zwei Geraden $G_i = p_i^{\mathsf{T}(G_i)} \in \mathfrak{G}_{01}$ ($G_i \in \mathfrak{G}_0$; $p_i \in \hat{P}_0$; $i = 1, 2$) parallel, wenn die G_i in $\mathfrak{p}_0$ zueinander parallel liegen.

(2) Aus den Mengen $\{p, q\}$ ($p, q \in \hat{P}_0$, $p \neq q$), falls p mit q durch keine Gerade aus $\mathfrak{G}_0 \cup \mathfrak{G}_{01}$ verbunden werden kann. Jedes so gewonnene $\{p, q\}$ soll mit p und q, aber keinem weiteren Punkt aus $\hat{P}_0$ inzidieren. Die Menge dieser Geraden bezeichnen wir mit $\mathfrak{G}_{02}$. Keine Gerade aus $\mathfrak{G}_{02}$ soll zu einer Geraden aus $\mathfrak{G}_0 \cup \mathfrak{G}_{01}$ parallel liegen. Dagegen seien zwei Geraden $\{p, q\}$ und $\{p', q'\}$ aus $\mathfrak{G}_{02}$ genau dann zueinander parallel, wenn es ein $\tau \in \mathsf{T}$ gibt derart, daß

die Punktmenge $\{p, q\}$ durch τ in die Menge $\{p', q'\}$ übergeführt wird: $\{p', q'\} = \{p, q\}^\tau = \{p^\tau, q^\tau\}$.

(3) Schließlich soll $\mathfrak{G}_1$ noch alle Paare (G, p) $[p \in \hat{P}_0;\ G \in \mathfrak{G}_0 \cup \mathfrak{G}_{02}]$ unter folgender Bedingung enthalten: In $\mathfrak{G}_0 \cup \mathfrak{G}_{02}$ existiert keine zu G parallele Gerade[9]) durch p (im Sinne unserer bisher eingeführten Parallelitätsdefinitionen). Die Menge der so erhaltenen Paare (G, p) nennen wir $\mathfrak{G}_{03}$. Jedes $(G, p) \in \mathfrak{G}_{03}$ soll mit p, aber keinem weiteren Punkt aus $\hat{P}_0$ inzidieren. Ferner sei (G, p) genau dann zu G' oder (G', p') parallel $(G, G' \in \mathfrak{G}_0 \cup \mathfrak{G}_{02};\ p, p' \in \hat{P}_0)$, wenn G zu G' parallel im Sinne unserer früheren Festsetzungen ist. Dagegen soll keine Gerade aus $\mathfrak{G}_{01}$ zu einer Geraden aus $\mathfrak{G}_{03}$ parallel sein.

Damit haben wir die Punkte, Geraden, Inzidenz- und Parallelitätsrelationen in $\mathfrak{p}_1$ vollständig festgelegt, also $\mathfrak{p}_1$ vollkommen bestimmt. Für die Geradenmenge von $\mathfrak{p}_1$ gilt

$$\mathfrak{G}_1 = \mathfrak{G}_0 \cup \mathfrak{G}_{01} \cup \mathfrak{G}_{02} \cup \mathfrak{G}_{03}. \tag{4.2}$$

Das in Fig. 1 gebrachte Schema ist in folgender Weise zu lesen: Ein Kreuz ($\times$) im Schnitt der Zeile $\mathfrak{G}_{0i}$ mit der Spalte $\mathfrak{G}_{0k}$ ($i, k = 0, 1, 2, 3$; $\mathfrak{G}_{00} = \mathfrak{G}_0$ gesetzt) besagt, daß es eine Gerade aus $\mathfrak{G}_{0i}$ gibt, die zu einer aus $\mathfrak{G}_{0k}$ parallel ist. Ein Kreis ($\circ$) in diesem Schnitt bedeutet das Gegenteil: Keine Gerade aus $\mathfrak{G}_{0i}$ ist zu einer aus $\mathfrak{G}_{0k}$ parallel.

	$\mathfrak{G}_0$	$\mathfrak{G}_{01}$	$\mathfrak{G}_{02}$	$\mathfrak{G}_{03}$
$\mathfrak{G}_0$	$\times$	$\times$	$\circ$	$\times$
$\mathfrak{G}_{01}$	$\times$	$\times$	$\circ$	$\circ$
$\mathfrak{G}_{02}$	$\circ$	$\circ$	$\times$	$\times$
$\mathfrak{G}_{03}$	$\times$	$\circ$	$\times$	$\times$

Fig. 1

3. Wir wollen jetzt zeigen, daß die soeben konstruierte Struktur $\mathfrak{p}_1$ eine t-Struktur darstellt, die eine translationstreue Erweiterung von $\mathfrak{p}_0$ ist. Wir zeigen zunächst:

HILFSSATZ 4.2. *$\mathfrak{p}_1$ ist eine a. P.*

BEWEIS. Nach Konstruktion der Geraden aus $\mathfrak{G}_1$ geht durch zwei verschiedene Punkte von $\mathfrak{p}_1$ genau eine Gerade, also gelten (a 0) und (a 1). Die in $\mathfrak{p}_1$ definierte Parallelenrelation erfüllt die drei Äquivalenzbedingungen: Reflexivität und Symmetrie sind unmittelbar ersichtlich, aber auch die Transitivität ist ohne weiteres nachzuprüfen, indem man alle möglichen Fälle durchdiskutiert. Zum Beweis von (a 2) ist zunächst zu sagen, daß die Existenz einer Parallele durch einen Punkt zu einer Geraden nach Konstruktion stets gesichert ist. Wir müssen noch die Eindeutigkeit zeigen, d. h. daß zwei verschiedene (im Sinne unserer in der vorigen Ziffer gemachten Festsetzungen) parallele Geraden $G_1, G_2 \in \mathfrak{G}_1$ sich niemals schneiden können. Dies ist klar, wenn beide Geraden in $\mathfrak{G}_0$ oder $\mathfrak{G}_{03}$ liegen, oder wenn die eine der Geraden in $\mathfrak{G}_0 \cup \mathfrak{G}_{01} \cup \mathfrak{G}_{02}$ und die andere in $\mathfrak{G}_{03}$ liegt. Sei nun $G_1 \in \mathfrak{G}_0$ und $G_2 \in \mathfrak{G}_{01}$. O.B.d.A. können wir $G_2 = p^{\mathsf{T}(G_1)}$ setzen. Aus $p^\tau \in G_1 \cap G_2$ für ein $\tau \in \mathsf{T}(G_1)$ würde $G_2 \subseteq G_1$ folgen, im Widerspruch zur Konstruktion von $\mathfrak{G}_{01}$. Sind $G_1, G_2 \in \mathfrak{G}_{01}$ und $G_1 \| G_2$, so gibt es ein $G \in \mathfrak{G}_0$ derart, daß $G_i = p_i^{\mathsf{T}(G)}$ ist $(i = 1, 2;\ p_i \in \hat{P}_0)$.

[9]) Ist $G \in \mathfrak{G}_{01}$ und $p \in \hat{P}_0$, so existiert bereits in $\mathfrak{G}_0 \cup \mathfrak{G}_{01}$ eine zu G parallele Gerade durch p, nämlich die durch die Punkte p^τ mit $\tau \in \mathsf{T}(G)$ gehende Gerade.

Aus $p \in G_1 \cap G_2$ folgt $p = p_1^{\tau_1} = p_2^{\tau_2}$ für geeignete $\tau_i \in \mathsf{T}(G)$, also $p_2 = p_1^{\tau_1 \tau_2^{-1}}$ und hieraus $G_1 = G_2$. Sei $G_i = \{p_i, q_i\} \in \mathfrak{G}_{02}$ $(i = 1, 2)$, $G_1 \| G_2$ und $G_1 \cap G_2 \neq \emptyset$. Dann ist o.B.d.A. $p_1 = p_2$. Aus der Parallelität von G_1 zu G_2 folgt $p_2 = p_1^{\tau}$, $q_2 = q_1^{\tau}$, oder $q_2 = p_1^{\tau}$, $p_2 = q_1^{\tau}$ für ein $\tau \in \mathsf{T}$. Im ersten Fall ist $\tau = \underline{1}$ (vgl. I, Satz 4.4), also $G_1 = G_2$; im zweiten Fall wäre $q_2 = p_2^{\tau}$, $\tau \neq \underline{1}$, mithin p_2 und q_2 durch eine Gerade aus $\mathfrak{G}_{01}$ verbunden, was nicht der Fall sein sollte. Damit ist auch (a 2) vollständig bewiesen.

Axiom (a 3) gilt für $\mathfrak{p}_0$ und damit erst recht für $\mathfrak{p}_1$, denn $\mathfrak{p}_1$ ist eine Erweiterung von $\mathfrak{p}_0$. Axiom (a 4) ist für die Geraden aus $\mathfrak{G}_0$ voraussetzungsgemäß und für die anderen nach Konstruktion erfüllt. Damit ist Hilfssatz 4.2 vollständig bewiesen.

4. In Ziffer 1 haben wir die Kollineationen γ von $\mathfrak{p}_0$ eindeutig auf $\hat{\mathfrak{p}}_0$ fortgesetzt. Nunmehr erweitern wir diese (eindeutig) in natürlicher Weise auf $\mathfrak{p}_1$ durch:

$$p^{\mathsf{T}(G)\gamma} = p^{\gamma\gamma^{-1}\mathsf{T}(G)\gamma} = p^{\gamma\,\mathsf{T}(G^{\gamma})} \qquad (G \in \mathfrak{G}), \tag{4.3}$$

$$\{p, q\}^{\gamma} = \{p^{\gamma}, q^{\gamma}\}, \tag{4.4}$$

$$(G, p)^{\gamma} = (G^{\gamma}, p^{\gamma}). \tag{4.5}$$

Wir müssen zeigen, daß die so fortgesetzten Abbildungen sinnvoll sind, d.h. daß die rechten Seiten von (4.3) bis (4.5) tatsächlich Geraden aus $\mathfrak{G}_{01}$ bzw. $\mathfrak{G}_{02}$ bzw. $\mathfrak{G}_{03}$ darstellen.

Zu (4.3). Gäbe es ein $G' \in \mathfrak{G}_0$ mit $p^{\mathsf{T}(G)\gamma} \subseteqq G'$, so wäre $p^{\mathsf{T}(G)} \subseteqq G'^{\gamma^{-1}} \in \mathfrak{G}_0$, im Widerspruch zur Voraussetzung. Also gilt $\mathfrak{G}_{01}^{\gamma} \subseteqq \mathfrak{G}_{01}$.

Zu (4.4). Würden p^{γ} und q^{γ} mit einem $G \in \mathfrak{G}_0 \cup \mathfrak{G}_{01}$ inzidieren, so wären $p, q \in G^{\gamma^{-1}} \in \mathfrak{G}_0 \cup \mathfrak{G}_{01}$. Daraus folgt $\mathfrak{G}_{02}^{\gamma} \subseteqq \mathfrak{G}_{02}$.

Zu (4.5). Gäbe es in $\mathfrak{G}_0 \cup \mathfrak{G}_{01} \cup \mathfrak{G}_{02}$ eine Gerade G', die mit p^{γ} inzidiert und zu G^{γ} parallel ist, so wäre $G'^{\gamma^{-1}} \in \mathfrak{G}_0 \cup \mathfrak{G}_{01} \cup \mathfrak{G}_{02}$ zu G parallel und würde mit p inzidieren; mithin ist $\mathfrak{G}_{03}^{\gamma} \subseteqq \mathfrak{G}_{03}$.

Insbesondere setzen sich auch die Translationen von $\mathfrak{p}_0$ als Kollineationen auf $\mathfrak{p}_1$ fort. Genauer gilt:

HILFSSATZ 4.3. *Die Fortsetzung einer Translation*[10]) *von* $\mathfrak{p}_0$ *auf* $\mathfrak{p}_1$ *ist eine zentrale Translation auf* $\mathfrak{p}_1$.

BEWEIS. Sei $\tau \neq \underline{1}$ eine Translation auf $\mathfrak{p}_0$ [deren Fortsetzung auf $\mathfrak{p}_1$ gemäß (4.3) bis (4.5) wir ebenfalls τ nennen]. Aus den Formeln (4.3) bis (4.5) und den Definitionen der Parallelität in $\mathfrak{p}_1$ folgt sofort, daß τ auch in $\mathfrak{p}_1$ jede Gerade in eine zu ihr parallele abbildet; in (4.3) ist dabei zu beachten, daß $\mathsf{T}(G) = \mathsf{T}(G^{\tau})$ $(G \in \mathfrak{G}_0)$ ist, denn $\mathfrak{p}_0$ ist voraussetzungsgemäß eine t-Struktur. Ferner läßt τ keinen Punkt aus $\hat{P}_0$ fest. Wäre nämlich $\{G, G'\}^{\tau} = \{G, G'\}$ für ein $\{G, G'\} \in \hat{P}_0 - P_0$ $(G, G' \in \mathfrak{G}_0)$, so würde $G^{\tau} = G$ und $G'^{\tau} = G'$ gelten wegen $G \nparallel G'$. Mithin hätte τ (als Translation von $\mathfrak{p}_0$ aufgefaßt) die voneinander ver-

[10]) Ebenso ist jede Fortsetzung einer Streckung von $\mathfrak{p}_0$ auf $\mathfrak{p}_1$ eine Streckung auf $\mathfrak{p}_1$. Der Beweis hierfür ist analog dem von Hilfssatz 4.3.

schiedenen uneigentlichen Punkte von G und G' zu Zentren, was wegen $\tau \neq \underline{1}$ nicht sein kann. Also ist τ auch auf $\mathfrak{p}_1$ eine Translation. (Ist $\tau = \underline{1}$ auf $\mathfrak{p}_0$, so auch auf $\mathfrak{p}_1$, was sofort aus I, Satz 4.1 folgt.) Es bleibt noch zu zeigen, daß τ auf $\mathfrak{p}_1$ zentral ist. O.B.d.A. nehmen wir wieder $\tau \neq \underline{1}$ an. Jedenfalls gibt es ein $G \in \mathfrak{G}_0$ derart, daß $G'^{\tau} = G'$ für alle $G' \in \mathfrak{G}_0$ gilt, die zu G parallel sind, denn τ ist auf $\mathfrak{p}_0$ zentral. Sei nun $G' \in \mathfrak{G}_1 - \mathfrak{G}_0$ und $G' \| G$, ferner p' ein auf G' gelegener Punkt von $\hat{P}_0$. In diesem Falle ist $G' = p'^{\mathsf{T}(G)} \in \mathfrak{G}_{01}$ und $G'^{\tau} = G'$. Somit bleibt auch jede in $\mathfrak{p}_1$ zu G parallele Gerade bei Anwendung von τ (als Ganzes) fest. Damit ist τ in $\mathfrak{p}_1$ als zentral erkannt.

HILFSSATZ 4.4. *Genau die nicht zu P_0 gehörenden Punkte aus $\hat{P}_0$ inzidieren mit Geraden aus $\mathfrak{G}_{03}$, die zu Geraden aus $\mathfrak{G}_0$ parallel sind.*

BEWEIS. (a) Sei $p \in P_0$. Jede durch p gehende Parallele zu einem $G \in \mathfrak{G}_0$ liegt ebenfalls in $\mathfrak{G}_0$, da ja $\mathfrak{p}_0$ als a. P. vorausgesetzt war. Durch ein solches p kann mithin keine Gerade aus $\mathfrak{G}_{03}$ gehen, die zu einer Geraden aus $\mathfrak{G}_0$ parallel ist.

(b) Sei jetzt $p \in \hat{P}_0 - P_0$. Dann ist $p = \{G_1, G_2\}$, wobei G_1 und G_2 (in $\mathfrak{p}_0$) windschiefe Geraden aus $\mathfrak{G}_0$ sind. Nach unserer zu Anfang dieses Paragraphen gemachten Voraussetzung (B) gibt es ein $G \in \mathfrak{G}_0$, welches weder zu G_1 noch zu G_2 parallel ist und für das $\mathsf{T}(G) = \{\underline{1}\}$ gilt. Die durch p gehende zu G parallele Gerade G^* von $\mathfrak{p}_1$ gehört nicht zu $\mathfrak{G}_0$, denn die einzigen durch p gehenden Geraden aus $\mathfrak{G}_0$ sind G_1 und G_2. Wegen $\mathsf{T}(G) = \{\underline{1}\}$ gehört sie auch nicht zu $\mathfrak{G}_{01}$. Da keine Gerade von $\mathfrak{G}_{02}$ zu einer aus $\mathfrak{G}_0$ parallel ist (vgl. Fig. 1) kann G^* auch nicht in $\mathfrak{G}_{02}$ liegen. Mithin gehört $G^* = (G, p)$ zu $\mathfrak{G}_{03}$.

HILFSSATZ 4.5. *Genau diejenigen Geraden aus $\mathfrak{p}_1$, die mit nur einem Punkt inzidieren, gehören zu $\mathfrak{G}_{03}$.*

BEWEIS. Jedenfalls inzidieren nach Konstruktion die Geraden aus $\mathfrak{G}_{03}$ mit nur einem Punkt. Dagegen enthält jede Gerade aus $\mathfrak{G}_{02}$ genau zwei und eine Gerade $p^{\mathsf{T}(G)} \in \mathfrak{G}_{01}$ genau $|\mathsf{T}(G)| > 1$ Punkte. Geraden aus $\mathfrak{G}_0$ können zwar in $\mathfrak{p}_0$ mit nur einem Punkt inzidieren, jedoch nicht in $\mathfrak{p}_1$. Enthalte nämlich ein $G \in \mathfrak{G}_0$ in $\mathfrak{p}_0$ nur einen Punkt p. Nach Satz 1.4 gibt es in $\mathfrak{p}_0$ ein nichtausgeartetes Viereck, dessen einer Eckpunkt p ist. Die anderen Eckpunkte mögen p_1, p_2 und p_3 sein. Wegen unserer Annahme hat keine der Verbindungsgeraden $\langle p_i, p_k \rangle_{\mathfrak{p}_0}$ $(i, k = 1, 2, 3;\ i \neq k)$ mit G in $\mathfrak{p}_0$ einen Schnittpunkt gemeinsam. Mindestens eine der Geraden, etwa $\langle p_1, p_2 \rangle$, ist zu G nicht parallel. Dann ist $\{G, \langle p_1, p_2 \rangle\} \in \hat{P}_0 = P_1$ ein zweiter Punkt, der mit G in $\mathfrak{p}_1$ inzidiert.

5. Sei nun γ^* eine *beliebige* Kollineation von $\mathfrak{p}_1$ [also nicht notwendig eine, die von einer Kollineation von $\mathfrak{p}_0$ induziert wird (im Sinne von (4.3) bis (4.5))]. Aus Hilfssatz 4.5 folgt wegen $|G^{\gamma^*}| = |G|$ für jedes $G \in \mathfrak{G}_1$ sofort $\mathfrak{G}_{03}^{\gamma^*} \subseteq \mathfrak{G}_{03}$. Da dasselbe für γ^{*-1} gilt, haben wir:

HILFSSATZ 4.6. *Für jede Kollineation γ^* von $\mathfrak{p}_1$ gilt*

$$\mathfrak{G}_{03}^{\gamma^*} = \mathfrak{G}_{03}. \tag{4.6}$$

Eine Translation τ^* von $\mathfrak{p}_1$ führt insbesondere jede Gerade in eine zu ihr parallele über. Zusammen mit den Hilfssätzen 4.4 und 4.6 ergibt diese Bemerkung $(\hat{P}_0 - P_0)^{\tau^*} = \hat{P}_0 - P_0$, also:

HILFSSATZ 4.7. *Jede Translation τ^* von $\mathfrak{p}_1$ bildet P_0 auf sich ab*[11]:

$$P_0^{\tau^*} = P_0. \tag{4.7}$$

Nach (4.7) ist die Bedingung (a') für eine translationstreue Erweiterung erfüllt (vgl. I, §4, Ziffer 5); aus Hilfssatz 4.3 folgt einerseits (b') und andererseits, daß jede Translation von $\mathfrak{p}_1$ zentral ist. Somit erhalten wir:

SATZ 4.1. *Die durch die freie translationstreue Erweiterung (im Sinne der in den Ziffern 1 und 2 angegebenen Konstruktion) der t-Struktur $\mathfrak{p}_0$ gewonnene Struktur $\mathfrak{p}_1$ ist eine über $\mathfrak{p}_0$ translationstreue t-Struktur.*

BEMERKUNG. Die Voraussetzungen (A) und (B) sind auch für $\mathfrak{p}_1$ gültig. Denn $\mathfrak{p}_1$ ist keine affine Ebene, weil $\mathfrak{p}_1$ Gerad n mit nur einen Punkt enthält (nämlich die Geraden aus $\mathfrak{G}_{03}$). Die nach Voraussetzung (B) in $\mathfrak{G}_0$ existierenden Geraden G_i $(i=1, \ldots, 5)$ mit $\mathsf{T}(G_i) = \{\underline{1}\}$ haben nach Satz 4.1 in $\mathfrak{p}_1$ dieselbe Eigenschaft, also ist auch (B) in $\mathfrak{p}_1$ erfüllt.

§5. Die freie translationstreue Erweiterungsebene $[\mathfrak{p}_0]$

1. Aus Satz 4.1 und der daran anschließenden Bemerkung folgt, daß für $\mathfrak{p}_1$ dieselben Voraussetzungen erfüllt sind wie für $\mathfrak{p}_0$. Fortgesetzte Anwendung desselben Konstruktionsverfahrens ergibt somit eine aufsteigende Kette von t-Strukturen

$$\mathfrak{p}_0 \leqq \mathfrak{p}_1 \leqq \mathfrak{p}_2 \leqq \cdots, \tag{5.1}$$

in der $\mathfrak{p}_{n+1}$ aus $\mathfrak{p}_n = (P_n, \mathfrak{G}_n)$ $(n = 0, 1, 2, \ldots)$ nach demselben Verfahren gebildet wird wie $\mathfrak{p}_1$ aus $\mathfrak{p}_0$ (s. §4). Die Punktmenge $\hat{P}_n$ entspricht dabei der Menge $\hat{P}_0$, die Geradenmengen $\mathfrak{G}_n$, $\mathfrak{G}_{n1}$, $\mathfrak{G}_{n2}$ und $\mathfrak{G}_{n3}$ den Mengen $\mathfrak{G}_0$, $\mathfrak{G}_{01}$, $\mathfrak{G}_{02}$ und $\mathfrak{G}_{03}$. Es ist also

$$P_{n+1} = \hat{P}_n, \qquad \mathfrak{G}_{n+1} = \mathfrak{G}_n \cup \mathfrak{G}_{n1} \cup \mathfrak{G}_{n2} \cup \mathfrak{G}_{n3}. \tag{5.2}$$

Wegen Satz 4.1 und I, Satz 4.7 ist jedes $\mathfrak{p}_n$ eine über $\mathfrak{p}_0$ translationstreue t-Struktur.

HILFSSATZ 5.1. *Die Vereinigung*

$$\mathfrak{A} = [\mathfrak{p}_0] = \bigcup_{i=0}^{\infty} \mathfrak{p}_i \tag{5.3}$$

(vgl. I, §2, Ziffer 5) *der t-Strukturen der Kette* (5.1) *ist eine affine Ebene.*

Hierzu müssen wir nur noch zeigen, daß zwei nichtparallele Geraden von $\mathfrak{A} = (P, \mathfrak{G})$ sich stets schneiden. Seien $G, G' \in \mathfrak{G}$, $G \nparallel G'$. Wegen $\mathfrak{G}_0 \subseteqq \mathfrak{G}_1 \subseteqq \cdots$ und $\mathfrak{G} = \bigcup_{i=0}^{\infty} \mathfrak{G}_i$ gibt es ein n derart, daß G und G' zu $\mathfrak{G}_n$ gehören. Haben

[11]) Satz und Beweis übertragen sich sofort auch auf Streckungen von $\mathfrak{p}_1$.

G und G' in $\mathfrak{p}_n$ noch keinen Punkt gemeinsam, so ist $\{G, G'\} \in P_{n+1}$ der gesuchte Schnittpunkt.

DEFINITION. Wir nennen die so erhaltene affine Ebene $\mathfrak{A} = [\mathfrak{p}_0]$ die *von $\mathfrak{p}_0$ erzeugte freie translationstreue Erweiterungsebene.*

2. Wir wollen jetzt zeigen, daß der Ausdruck „translationstreu" bei der Benennung von $\mathfrak{A}$ zu Recht besteht, d.h. daß $\mathfrak{A}$ translationstreu über $\mathfrak{p}_0$ ist. Die Bedingung (b') [vgl. I, §4, Ziffer 5] ist sofort einzusehen. Sei nämlich τ eine Translation von $\mathfrak{p}_0$. Nach Hilfssatz 4.3 läßt sich τ eindeutig auf $\mathfrak{p}_1$, also auch auf jedes $\mathfrak{p}_n$ fortsetzen. (Diese Fortsetzung können und wollen wir ebenfalls τ nennen.) Da jeder Punkt und jede Gerade von $\mathfrak{A}$ einem P_n bzw.

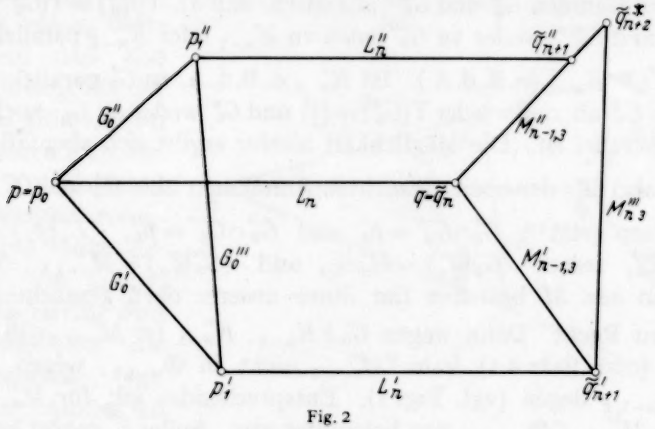

Fig. 2

$\mathfrak{G}_n$ angehört, ist die Abbildung τ auf der ganzen Ebene definiert. Sie ist dort eine Translation, da sie es auf jedem $\mathfrak{p}_n$ ist.

Es bleibt somit noch (a') zu beweisen:

HILFSSATZ 5.2. *Für jede Translation τ von $\mathfrak{A}$ gilt $P_0^{\tau} = P_0$.*

Wir müssen also zeigen, daß es in $\mathfrak{A}$ keine Translation geben kann, die ein $p \in P_0$ in ein $q \in P_n - P_{n-1}$ mit $n \geqq 1$ abbildet. Dieser Beweis wird so geführt: Es wird eine kleine Desarguessche Konfiguration mit p, q als ein Paar entsprechender Dreieckspunkte angegeben, die sich nicht schließt (vgl. Fig. 2). Hieraus folgt dann sofort, daß keine Translation p in q überführen kann.

Zur Abkürzung setzen wir $P_n - P_{n-1} = \tilde{P}_n$ und $\mathfrak{G}_n - \mathfrak{G}_{n-1} = \tilde{\mathfrak{G}}_n$ $(n = 1, 2, \ldots)$. Ferner verabreden wir der einfacheren Schreibweise wegen folgendes: Ein an einem Punkt p oder einer Geraden G angebrachter Index i $(= 0, 1, \ldots)$ besagt, daß das betreffende Element in P_i bzw. $\mathfrak{G}_i$ liegt: $p_i \in P_i$, $G_i \in \mathfrak{G}_i$. Aufsetzen einer Tilde $(\sim)$ auf p_i bedeutet, daß der betreffende Punkt zu $\tilde{P}_i$ gehört; $\tilde{p}_i \in \tilde{P}_i$. Ein an einer Geraden angebrachter Doppelindex ik $(k = 1, 2, 3)$ besagt, daß sie in $\mathfrak{G}_{ik}$ liegt: $G_{ik} \in \mathfrak{G}_{ik}$.

Seien nun $p = p_0$ und $q = \tilde{q}_n$ $(n \geqq 1)$ im Sinne der eben gemachten Verabredung beliebig vorgegeben; wir setzen $\langle p_0, \tilde{q}_n \rangle = L_n$ (s. auch Fig. 2). Durch p_0 legen wir zwei von L_n verschiedene Geraden G_0' und G_0'' mit $\mathsf{T}(G_0') = \mathsf{T}(G_0'') = \{\underline{1}\}$,

ferner sei G_0''' eine nicht durch p_0 gehende und nicht zu G_0', G_0'' und L_n parallele Gerade, ebenfalls mit $\mathsf{T}(G_0''')=\{\underline{1}\}$. Nach unserer Voraussetzung (B) und wegen der Gültigkeit von (a 3) in p_0 kann man solche Geraden immer finden. Es seien K_{n-1} und K'_{n-1} die $\tilde{q}_n$ erzeugenden Geraden, d.h. $\tilde{q}_n=\{K_{n-1}, K'_{n-1}\}$. Für G_0' und G_0'' fordern wir nun außerdem noch, daß sie weder zu K_{n-1} noch zu K'_{n-1} parallel sind. Um zu beweisen, daß diese Forderung mit der vorher genannten verträglich ist, machen wir die folgende Fallunterscheidung:

(a) L_n ist von K_{n-1} und K'_{n-1} verschieden. Dann ist $L_n\in\tilde{\mathfrak{G}}_n$, weil durch $\tilde{q}_n$ aus $\mathfrak{G}_{n-1}$ nur die beiden $\tilde{q}_n$ erzeugenden Geraden gehen. Nehmen wir an, G_0' und G_0'' seien zu K_{n-1} und K'_{n-1} parallel; nach (B) können wir dann G_0' und G_0'' so zu Geraden G_0^* und G_0^{**} abändern, daß 1). $\mathsf{T}(G_0^*)=\mathsf{T}(G_0^{**})=\{\underline{1}\}$ ist und 2). G_0^* und G_0^{**} weder zu G_0''' noch zu K_{n-1} oder K'_{n-1} parallel sind.

(b) Sei $L_n=K_{n-1}$ (o.B.d.A.). Ist K'_{n-1} o.B.d.A. zu G_0' parallel, so ändern wir G_0' so zu G_0^* ab, daß wieder $\mathsf{T}(G_0^*)=\{\underline{1}\}$ und G_0^* weder zu G_0''' noch zu K_{n-1} oder K'_{n-1} parallel ist. Die Möglichkeit hierfür ergibt sich ebenfalls aus (B).

Es sind also alle der oben gemachten Annahmen über G_0' und G_0'' erfüllbar.

Wir setzen jetzt[12]) $G_0'\cap G_0'''=p_1'$ und $G_0''\cap G_0'''=p_1''$, $\langle L_n\|p_1'\rangle=L_n'$ und $\langle L_n\|p_1''\rangle=L_n''$, ferner $\langle G_0'\|\tilde{q}_n\rangle=M'_{n-1,3}$ und $\langle G_0''\|\tilde{q}_n\rangle=M''_{n-1,3}$. Die Indizierungen an den M bestehen (im Sinne unserer oben gemachten Vereinbarungen) zu Recht. Denn wegen $G_0'\nparallel K_{n-1}$, K'_{n-1} ist $M'_{n-1,3}\in\tilde{\mathfrak{G}}_n$. Wegen $\mathsf{T}(G_0')=\{\underline{1}\}$ (und Satz 4.1) kann $M'_{n-1,3}$ nicht in $\mathfrak{G}_{n-1,1}$, wegen $M'_{n-1,3}\|G_0'$ nicht in $\mathfrak{G}_{n-1,2}$ liegen (vgl. Fig. 1). Entsprechendes gilt für $M''_{n-1,3}$. Also gilt $M'_{n-1,3}$, $M''_{n-1,3}\in\mathfrak{G}_{n-1,3}$, was behauptet war. Außer $\tilde{q}_n$ gehört kein Punkt von $M'_{n-1,3}$ bzw. $M''_{n-1,3}$ zu P_n. Mithin ist $M'_{n-1,3}\cap L_n'=\tilde{q}'_{n+1}$ und $M''_{n-1,3}\cap L_n''=\tilde{q}''_{n+1}$. Offensichtlich sind $M'_{n-1,3}$ und L_n' die erzeugenden Geraden von $\tilde{q}'_{n+1}$, d.h. $\tilde{q}'_{n+1}=\{M'_{n-1,3}, L_n'\}$. Aus denselben Gründen wie eben ist $\langle G_0'''\|q'_{n+1}\rangle=M'''_{n,3}\in\mathfrak{G}_{n,3}$. Da kein weiterer Punkt von $M'''_{n,3}$ in P_{n+1} liegt, gilt $M'''_{n3}\cap L_n''=\tilde{q}_{n+2}\neq\tilde{q}''_{n+1}$.

Gäbe es nun in $\mathscr{A}=[p_0]$ eine Translation τ, die p_0 in $\tilde{q}_n$ überführt, so wäre $p_1''^{\tau}=\tilde{q}''_{n+1}=\tilde{q}_{n+2}$, im Widerspruch zu dem eben gezeigten. Somit ist $P_0^{\tau}\subseteq P_0$, ebenso gilt aber auch $P_0^{\tau^{-1}}\subseteq P_0$, also Hilfssatz 5.2.

3. Damit ist gezeigt, daß $\mathscr{A}$ translationstreu über p_0 ist:

SATZ 5.1. *Die von einer t-Struktur p_0 erzeugte freie translationstreue Erweiterungsebene $[p_0]$ ist translationstreu*[13]) *über p_0.*

Hiermit ist das im vorletzten Paragraphen gestellte Problem, eine t-Struktur translationstreu in eine affine Ebene einzubetten, vollständig gelöst.

Insbesondere gilt die folgende Spezialisierung:

ZUSATZ 1. *Jede zentrale Translationsstruktur ist translationstreu in eine affine Ebene einbettbar.*

---

[12]) Vgl. Fußnote 6.

[13]) Jedenfalls unter der Voraussetzung (B) von § 4, die aber, wie dort gezeigt, keine Beschränkung der Allgemeinheit darstellt.

Aus den Ergebnissen von II (insbesondere Satz 1.3) folgt nun:

ZUSATZ 2. *Jede abelsche Gruppe mit einer Partition ist Translationsgruppe einer affinen Ebene.*

§ 6. Die Streckungen der freien translationstreuen Erweiterungsebene

1. Durch die Untersuchungen des vorigen Paragraphen haben wir eine vollständige Übersicht über die Translationen der freien translationstreuen Erweiterungsebene $\mathfrak{A} = [\mathfrak{p}_0]$ von $\mathfrak{p}_0$ erhalten: Sie stimmen mit denen von $\mathfrak{p}_0$ überein. Wir zeigen jetzt weitergehend, daß alle affin-axialen Kollineationen (vgl. I, § 4, Ziffer 4) von $\mathfrak{A}$ mit denen von $\mathfrak{p}_0$ übereinstimmen, daß also $\mathfrak{A}$ eine hinsichtlich der Gruppe der affin-axialen Kollineationen (d.h. der von den Translationen und Streckungen erzeugten Gruppe) treue Erweiterung von $\mathfrak{p}_0$ ist. Der Beweis hierfür wird ähnlich dem für die Translationstreue (vgl. § 5) geführt. Analog zu Hilfssatz 5.2 zeigen wir zunächst:

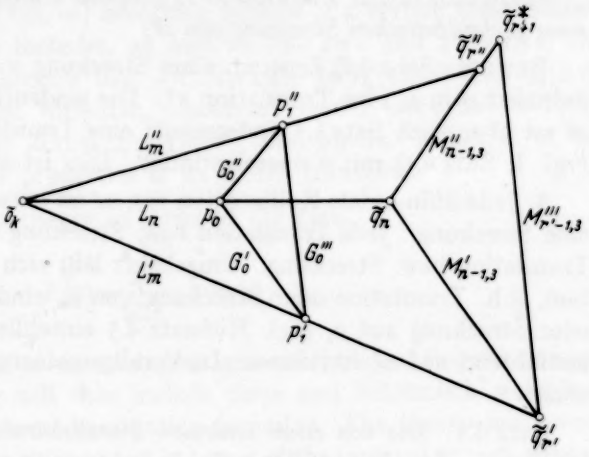

Fig. 3

HILFSSATZ 6.1. *Für jede Streckung σ von $\mathfrak{A}$ gilt $P_0^\sigma = P_0$.*

Zum Beweis machen wir wieder dieselben Verabredungen über die Indizierung von Punkten und Geraden wie im vorigen Paragraphen. Sei $\tilde{o}_k$ das Zentrum der Streckung σ von $\mathfrak{A}$, von der wir annehmen wollen, daß sie $\mathfrak{p}_0$ in $\tilde{q}_n = \{K_{n-1}, K'_{n-1}\}$ mit $n \geqq 1$ abbilde. (Vgl. auch Fig. 3.) Die Punkte $\tilde{o}_k$, $\mathfrak{p}_0$ und $\tilde{q}_n$ liegen mithin auf einer Geraden L_n. Es seien wieder G'_0, G''_0 und G'''_0 drei nicht zueinander parallele Geraden, für die $\mathsf{T}(G'_0) = \mathsf{T}(G''_0) = \mathsf{T}(G'''_0) = \{1\}$, $\mathfrak{p}_0 \in G'_0, G''_0$, $\mathfrak{p}_0 \notin G'''_0$ und $K_{n-1}, K'_{n-1} \nparallel G'_0, G''_0$ gelten. Die Existenz solcher Geraden [unter Voraussetzung (B) von § 4] ergibt sich ebenso wie beim Beweis von Hilfssatz 5.2. Außerdem müssen wir noch $\tilde{o}_k \notin G'''_0$ verlangen. Auch dies ist noch zu erreichen, denn aus Satz 1.4 zusammen mit Axiom (a 2) folgt sofort, daß es in einer t-Struktur zu jeder Geraden mindestens zwei von ihr verschiedene und zu ihr parallele Geraden gibt.

Wir setzen jetzt $\mathfrak{p}'_1 = G'_0 \cap G'''_0$, $\mathfrak{p}''_1 = G''_0 \cap G'''_0$, $L'_m = \langle \tilde{o}_k, \mathfrak{p}'_1 \rangle$, $L''_m = \langle \tilde{o}_k, \mathfrak{p}''_1 \rangle$ [wobei $m = \mathrm{Max}(k, 1)$ gesetzt wurde; aus $\tilde{o}_k \notin G'''_0$ folgt $L'_m \neq L''_m$], ferner $\langle G'_0 \| \tilde{q}_n \rangle = M'_{n-1,3}$ und $\langle G''_0 \| \tilde{q}_n \rangle = M''_{n-1,3}$; die Richtigkeit der Indizierung im Sinne unserer Verabredung ergibt sich ebenso wie die entsprechende im vorigen Paragraphen. Außer $\tilde{q}_n$ enthalten $M'_{n-1,3}$ und $M''_{n-1,3}$ keinen Punkt aus P_n. Mithin ist $L'_m \cap M'_{n-1,3} = \tilde{q}'_{r'}$ und $L''_m \cap M''_{n-1,3} = \tilde{q}''_{r''}$, wobei (o. B. d. A.) $n < r'' \leqq r'$

ist. Wegen dieser Ungleichung ist $M'_{n-1,3}$ eine der beiden erzeugenden Geraden von $\tilde{q}'_{r'}$. Sei nun $\tilde{q}'_{r'}=\{M'_{n-1,3}, S_t\}$ $(t<r')$. Ist zufällig $S_t \| G'''_0$, so ändern wir G'''_0 so ab, daß es nicht mehr zu S_t parallel ist, ohne die anderen Eigenschaften von G'''_0 zu verletzen; dies ist wegen (B) möglich. O.B.d.A. sei also $S_t \nparallel G'''_0$. Dann ist $\langle G'''_0 \| \tilde{q}'_{r'}\rangle = M'''_{r'-1,3}$ mit richtiger Indizierung. Außer $\tilde{q}'_{r'}$ liegt kein Punkt von $M'''_{r'-1,3}$ in $P_{r'}$, also ist $M'''_{r'-1,3} \cap M''_{n-1,3} = \tilde{q}_{r'+1} \neq \tilde{q}''_{r''}$, letzteres wegen $r'' < r'+1$. Aus $p_0^{\sigma}=\tilde{q}_n$ für eine Streckung σ von $\mathfrak{A}$ würde aber $p_1^{''\sigma}=\tilde{q}''_{r''}=\tilde{q}_{r'+1}$ folgen, im Widerspruch zu dem eben gezeigten. Also gilt $P_0^{\sigma} \subseteqq P_0$, ebenso $P_0^{\sigma^{-1}} \subseteqq P_0$, d.h. Hilfssatz 6.1.

2. HILFSSATZ 6.2. *Ein nicht in P_0 gelegener Punkt von $\mathfrak{A}$ ist niemals Zentrum einer nichtidentischen Streckung von $\mathfrak{A}$.*

BEWEIS. Sei $o \notin P_0$ Zentrum einer Streckung σ von $\mathfrak{A}$. Nach Hilfssatz 6.1 induziert σ in $\mathfrak{p}_0$ eine Translation σ^*. Die eindeutige Fortsetzung von σ^* auf $\mathfrak{A}$ ist aber nach Satz 5.1 andererseits eine Translation auf $\mathfrak{A}$, die natürlich (vgl. I, Satz 4.1) mit σ übereinstimmt. Dies ist aber nur für $\sigma = \underline{1}$ möglich.

3. Jede affin-axiale Kollineation von $\mathfrak{A}$ ist entweder eine Translation oder eine Streckung. Jede Translation bzw. Streckung von $\mathfrak{A}$ induziert in $\mathfrak{p}_0$ eine Translation bzw. Streckung. Umgekehrt läßt sich jede affin-axiale Kollineation, d.h. Translation oder Streckung von $\mathfrak{p}_0$ eindeutig zu einer Translation oder Streckung auf $\mathfrak{p}_1$ (vgl. Hilfssatz 4.3 einschließlich Fußnote) und somit fortfahrend auf $\mathfrak{A}$ fortsetzen. In Verallgemeinerung von Satz 5.1 gilt demnach:

SATZ 6.1. *Die von einer zentralen Parallelstruktur $\mathfrak{p}_0$ erzeugte freie translationstreue Erweiterungsebene $\mathfrak{A}=[\mathfrak{p}_0]$ ist eine treue Erweiterung hinsichtlich der Gruppe der affin-axialen Kollineationen, d.h.: Jede affin-axiale Kollineation von $\mathfrak{A}$ induziert in $\mathfrak{p}_0$ eine affin-axiale Kollineation, und umgekehrt gibt es zu jeder affin-axialen Kollineation π von $\mathfrak{p}_0$ eindeutig eine affin-axiale Kollineation von $\mathfrak{A}$, die π in $\mathfrak{p}_0$ induziert.*

Bei dieser eineindeutigen Zuordnung zwischen den affin-axialen Kollineationen von $\mathfrak{A}$ und denen von $\mathfrak{p}_0$ entsprechen sich jeweils die Translationen und die Streckungen. Insbesondere ist kein außerhalb der Punktmenge P_0 von $\mathfrak{p}_0$ gelegener Punkt Zentrum einer nichtidentischen Streckung.

Literatur

ANDRÉ, J.: Über Perspektivitäten in endlichen affinen Ebenen. Arch. d. Math. **9**, 228–235 (1958, a).

– Sur quelques plans euclidiens avec un groupe de translations. Convegno Internazionale „Reticoli e geometrie proiettive", Palermo-Messina, p. 1–2 (1958, b).

HALL, M.: Projective planes. Trans. Amer. Math. Soc. **54**, 229–277 (1943).

Braunschweig, Mathematisches Institut der Technischen Hochschule

(Eingegangen am 24. Oktober 1960)

MILLER, J. B.
Math. Zeitschr. 76, 257–269 (1961)

A class of series involving integral kernels. I

By
J. B. MILLER

1. Introduction

One particular generalization of Fourier sine and cosine integral transforms is to transformations in $L^2(0, \infty)$ involving kernels of a certain class defined by G. N. WATSON, which includes, as well as $2\sin 2\pi x$ and $2\cos 2\pi x$, the Hankel kernel $x^{\frac{1}{2}} J_\nu(x)$, kernels like $4K_0(4\pi x^{\frac{1}{2}}) - 2\pi Y_0(4\pi x^{\frac{1}{2}})$ and $J_0(2x^{\frac{1}{2}})$ which arise in summation formulae, associated kernel pairs like $x^{\frac{1}{2}} Y_\nu(x)$, $x^{\frac{1}{2}} \boldsymbol{H}_\nu(x)$, the Hilbert-type kernel $2/\pi(1-x^2)$, and others. A kernel of the class is defined by simple properties of its Mellin transform.

In this paper I give a theory of series of the type

$$F(x) = \sum_{n=1}^{\infty} a_n K(n x), \tag{1.1}$$

where K is such a kernel, as an analogous generalization of Fourier sine and cosine series; the theory will thus include these and Schlömilch series [1]) as special cases, as well as some summation formulae. The theory has several distinctive features. First, it is stated in terms of the interval $(0, \infty)$, whereas Fourier and Schlömilch series are for "arbitrary" functions defined over a finite interval only. Thus secondly, it follows intuitively that expansion of an arbitrary function $f(x)$ in the form (1.1) is not necessarily possible; f must be periodic, or have some equivalent property. Thirdly, convergence of the series is generally not possible in $L^2(0, \infty)$, the space in which the integral transforms are defined; it seems necessary instead to work in a certain Hilbert space $\mathscr{H}_{-1}$ which contains $L^2(0, \infty)$, and whose elements (called "sequence-functions") are generalized functions with some of the properties of distributions. Those sequence-functions which possess expansions are related to Fourier series through their Mellin transforms.

Among the examples which arise most simply from the theory are a number of summation formulae of the type considered by OPPENHEIM [2]), associated with Dirichlet series. This is because (1.1) corresponds to a relation

$$\mathfrak{F}(s) = \mathfrak{K}(s) \sum_{n=1}^{\infty} a_n n^{-s} \tag{1.2}$$

between the Mellin transforms of F and K.

---

[1]) References to the literature are given in [2], §7.10.3.
[2]) [6]. See also [2], §17.11.

The paper begins with a summary (§2) of the definition and properties of $\mathscr{H}_{-1}$, which are discussed in greater detail in [5]. §3 deals with the relation between sequence-function and coefficients, §4 with the convergence of the series. In §5 the class of sequence-functions possessing expansions is characterized. In §6 we draw an analogy with Fourier series, showing that the K-transform of a K-series is a sum of delta sequence-functions. Some examples are given in §7.

2. Notation and preliminary results

We write $f^{(-1)}(x)=\int_0^x f(t)\,dt$, and say that $f\in\mathscr{G}_{-1}$ if $x^{-1}f^{(-1)}(x)\in L^2(0,\infty)\equiv L^2$. We write

$$(f,g)_{-1}=\int_0^\infty t^{-2}f^{(-1)}(t)\,\bar{g}^{(-1)}(t)\,dt,\qquad \|f\|_{-1}=\sqrt{(f,f)_{-1}}$$

for the scalar product and norm in $\mathscr{G}_{-1}$, which is an incomplete space. We denote by $\mathscr{H}_{-1}$ the Hilbert space which is the completion in norm of $\mathscr{G}_{-1}$; i.e. an element F of $\mathscr{H}_{-1}$ is a "sequence-function", definable by a class of equivalent Cauchy sequences (f_n), $f_n\in\mathscr{G}_{-1}$ (or even $f_n\in L^2$, since L^2 is dense in $\mathscr{G}_{-1}$) for $n=1,2,\ldots$, for which $\|f_n-f_m\|_{-1}\to 0$ as $\min(n,m)\to\infty$. Then $L^2<\mathscr{G}_{-1}<\mathscr{H}_{-1}$. These spaces are described in [5]. We write $\lim_{(-1)}$ for a limit in $\mathscr{H}_{-1}$ norm.

The Mellin transform of $f(x)$ is $\mathfrak{F}(s)$, $s=\frac{1}{2}+it$, where

$$\mathfrak{F}(s)=\int_0^\infty f(x)\,x^{s-1}\,dx,\qquad f(x)=\frac{1}{2\pi i}\int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty}\mathfrak{F}(s)\,x^{-s}\,ds.$$

If $f\in L^2(0,\infty)$, $\mathfrak{F}(\frac{1}{2}+it)\in L^2(-\infty,\infty)$; the integrals then converge in mean square. The Parseval relation for Mellin transforms, is

$$\int_0^\infty f(x)\,g(x)\,dx=\frac{1}{2\pi i}\int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty}\mathfrak{F}(s)\,\mathfrak{G}(1-s)\,ds, \tag{2.1}$$

and implies

$$\int_0^\infty\{f(x)\}^2\,dx=\frac{1}{2\pi}\int_{-\infty}^{\infty}|\mathfrak{F}(\tfrac{1}{2}+it)|^2\,dt, \tag{2.2}$$

provided f is real. The Mellin transform of

$$x^{-1}f^{(-1)}(x)\quad\text{ is }\quad\mathfrak{F}_{-1}(s)=(1-s)^{-1}\mathfrak{F}(s).$$

These transforms are described in [7], pp. 94–5.

A Watson kernel[3]) is a sequence-function, K say, for which is defined indirectly a Mellin transform $\mathfrak{K}$ by

$$(1-s)^{-1}\mathfrak{K}(s)=\operatorname*{l.i.m.}_{X\to\infty}\int_{1/X}^{X}x^{-1}K^{(-1)}(x)\cdot x^{s-1}\,dx, \tag{2.3}$$

[3]) See [7], Ch. VIII. Below we use simply the designation "kernel".

with the property

$$1/A^* \leqq |\mathfrak{K}(\tfrac{1}{2}+it)| \leqq A \qquad (-\infty < t < \infty) \tag{2.4}$$

for some positive constants A, A^*. There will then exist an associated kernel K^*, determined by its Mellin transform $\mathfrak{K}^*$ by

$$\mathfrak{K}(\tfrac{1}{2}+it)\,\mathfrak{K}^*(\tfrac{1}{2}-it) = 1. \tag{2.5}$$

The kernel pair gives rise to a reciprocity

$$g(x) = \int_0^\infty K(xt)\,f(t)\,dt, \qquad f(x) = \int_0^\infty K^*(xt)\,g(t)\,dt.$$

The precise sense in which these transform relations exist is discussed in [5], §11. Some important cases are included if furthermore we suppose

$$|\mathfrak{K}(\tfrac{1}{2}+it)| = |\mathfrak{K}^*(\tfrac{1}{2}+it)| = 1, \tag{2.6}$$

so that K and K^* are complex conjugates. It is clear from the properties of Mellin transforms that (2.4) implies the convergence in L^2 of the integral in (2.3), so that necessarily $K \in \mathscr{H}_{-1}$.

A Watson kernel satisfies the integral formula

$$\int_0^\infty (nx)^{-1} K^{(-1)}(nx)\cdot(mx)^{-1} K^{*(-1)}(mx)\,dx = [\max(n,m)]^{-1}; \tag{2.7}$$

for by (2.1) and (2.5) the lefthand side equals

$$\frac{1}{2\pi i}\int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} \frac{n^{-s}\,\mathfrak{K}(s)}{1-s}\,\frac{m^{s-1}\,\mathfrak{K}^*(1-s)}{s}\,ds = \frac{1}{m}\,\frac{1}{2\pi i}\int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} \frac{(n/m)^{-s}}{s(1-s)}\,ds$$

which equals the righthand side. If we abbreviate $K(mx)$ to K_m, (2.7) becomes

$$(K_n, \bar{K}_m^*)_{-1} = \frac{1}{m}\,\frac{1}{2\pi i}\int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} \frac{(n/m)^{-s}}{s(1-s)}\,ds = [\max(n,m)]^{-1}. \tag{2.8}$$

When (2.6) holds, we can replace $\bar{K}_m^*$ here by K_m.

3. The series $\sum a_n K_n$; formulae for the ceofficients

Let K be a given kernel of $\mathscr{H}_{-1}$. Suppose that the sequence of numbers $\{a_n\}$ is such that as $N \to \infty$ the sequence-function

$$P_N(x) = \sum_{n=1}^{N} a_n K(nx) \tag{3.1}$$

converges in $\mathscr{H}_{-1}$, to P say. The sum P is uniquely determined by the coefficients and the kernel, by the definition of $\mathscr{H}_{-1}$. Conversely, if P and the kernel are known, the coefficients can be determined, in the following manner.

Since the scalar product in $\mathscr{H}_{-1}$ is continuous,

$$(3.2)\qquad \begin{cases} (P, \bar{K}_m^*)_{-1} = \lim\limits_{N\to\infty} (P_N, \bar{K}_m^*)_{-1} \\ \qquad = \lim\limits_{N\to\infty} \sum\limits_{n=1}^{N} a_n (K_n, \bar{K}_m^*)_{-1} = \sum\limits_{n=1}^{\infty} a_n [\max(n, m)]^{-1}. \end{cases}$$

Let

$$\sigma_m = (P, \bar{K}_m^*)_{-1} \qquad (m \geqq 1), \qquad \sigma_0 = 0.$$

Then

$$\sigma_m = \frac{1}{m} \sum_{n=1}^{m} a_n + \sum_{n=m+1}^{\infty} \frac{a_n}{n}$$

whence

$$\sigma_{m-1} - \sigma_m = \frac{1}{m(m-1)} \sum_{n=1}^{m-1} a_n \qquad (m \geqq 2),$$

and

$$(3.3)\qquad \frac{a_m}{m} = 2m\,\sigma_m - (m-1)\,\sigma_{m-1} - (m+1)\,\sigma_{m+1} \qquad (m \geqq 1).$$

These equations suggest some alternate formal procedures. If the expression (3.3) is substituted in $P = \sum a_n K_n$ and a formal rearrangement made in the series, we find

$$P = \sum m\,\sigma_m \{2m\,K_m - (m+1)\,K_{m+1} - (m-1)\,K_{m-1}\}.$$

If fact, if

$$L_m = m\{2m\,\bar{K}_m^* - (m+1)\,\bar{K}_{m+1}^* - (m-1)\,\bar{K}_{m-1}^*\} \quad (m \geqq 2),$$
$$L_1 = 2(\bar{K}_1^* - \bar{K}_2^*),$$

it is easily verified from (2.8) that

$$(K_n, L_m)_{-1} = 1 \quad (n = m), \quad 0 \quad (n \neq m) \qquad (m, n \geqq 1),$$

i.e. that $\{K_n\}$, $\{L_n\}$ form a pair of binormal sequences in $\mathscr{H}_{-1}$, and the preceding calculation takes the form

$$(P, L_m)_{-1} = (\sum a_n K_n, L_m)_{-1} = \sum a_n (K_n, L_m)_{-1} = a_m.$$

A third possible procedure if $K = \bar{K}^*$ is to show from (2.8) that $\{M_n\}$ where

$$M_m = m\,K_m - (m-1)\,K_{m-1} \qquad (m \geqq 2), \qquad M_1 = K_1$$

forms an orthonormal set in $\mathscr{H}_{-1}$, and from an expansion $P = \sum b_n M_n$ derive $b_m = (P, M_m)_{-1}$. We shall confine our attention to the form of series (3.1).

We note the Parseval-type formula

$$(3.4)\qquad (P, \bar{P}^*)_{-1} = \sum_{n=1}^{\infty} a_n \sigma_n,$$

whose proof is a simple consequence of the continuity of the scalar product:

$$(P, \bar{P}^*)_{-1} = \lim_{N\to\infty} \lim_{M\to\infty} \sum_{n=1}^{N} \sum_{m=1}^{M} \frac{a_n a_m}{\max(n, m)} = \lim_{N\to\infty} \sum_{n=1}^{N} a_n \sigma_n.$$

4. Convergence of the series

The convergence of the series $\sum a_n K_n$ *in* $\mathscr{H}_{-1}$ *is a property of the sequence* $\{a_n\}$, *and does not depend upon the kernel, provided* (2.6) *holds.* For let P_N be the partial sum (3.1); it converges to a limit in $\mathscr{H}_{-1}$ if and only if $\{P_N\}$ is a Cauchy sequence, and for $M<N$ we have

$$(4.1)\qquad \begin{cases} \|P_N - P_M\|_{-1}^2 = \left(\sum_{M+1}^{N} a_n K_n, \sum_{M+1}^{N} a_m K_m\right)_{-1} \\ \qquad = \sum_{M+1}^{N}\sum_{M+1}^{N} \dfrac{a_n \bar{a}_m}{\max(n, m)} \\ \qquad = \sum_{n=M+1}^{N} \left\{ \dfrac{a_n}{n} \sum_{m=M+1}^{n} \bar{a}_m + \dfrac{\bar{a}_n}{n} \sum_{m=M+1}^{n} a_m - \dfrac{|a_n|^2}{n} \right\}. \end{cases}$$

We shall call a sequence $\{a_n\}$ for which this sum approaches zero as $M \to \infty$, uniformly with N $(>M)$ a "w-sequence"; *if* (2.6) *holds, the series* $\sum a_n K_n$ *converges in* $\mathscr{H}_{-1}$ *if and only if* $\{a_n\}$ *is a* w*-sequence.*

It is easily seen, by writing $1 \leqq n^{\frac{1}{2}} m^{-\frac{1}{2}}$ in the inside sums of (4.1) (with moduli), that

$$(4.2)\qquad \sum_{n=1}^{\infty} |a_n|\, n^{-\frac{1}{2}} < \infty$$

is a sufficient condition for $\{a_n\}$ to be a w-sequence. This condition may be improved a little, as we now show.

THEOREM 1. *A set of sufficient conditions for* $\{a_n\}$ *to be a* w*-sequence is:*

(i) *the abscissa of absolute convergence* $\bar{\sigma}$ *of the Dirichlet series*

$$(4.3)\qquad \alpha(s) = \sum_{n=1}^{\infty} a_n n^{-s} \qquad (s = \sigma + it)$$

is less than 1;

(ii) *there exists a real number* c, $\bar{\sigma} < c < 1$, *such that* $\alpha(s)$ *may be continued analytically into the region* $\sigma \geqq 1 - c$, *the function* $\alpha(s)$ *having no singularities in the finite part of the strip* $1 - c \leqq \sigma \leqq c$;

(iii) $\alpha(\sigma + it) = o\left(|t|^2\right)$ *as* $t \to \pm\infty$, *uniformly for* $1 - c \leqq \sigma \leqq c$, *and* $\alpha(1 - c + it) = O\left(|t|^{1-\varepsilon}\right)$ *as* $t \to \pm\infty$, *for some positive* ε.

PROOF. From the previous discussion, we may assume $\bar{\sigma} > \frac{1}{2}$. Write $\bar{\alpha}(s)$ for $\sum \bar{a}_n n^{-s}$, and

$$S_{N,M} = \sum_{n=1}^{N}\sum_{m=1}^{M} a_n \bar{a}_m [\max(n, m)]^{-1}.$$

It will be sufficient if $S_{N,M}$ tends to a limit as $\min(N, M) \to \infty$. Now from (2.8),

$$(4.4)\qquad \begin{cases} S_{N,M} = \sum_{1}^{N}\sum_{1}^{M} \dfrac{a_n \bar{a}_m}{m} \dfrac{1}{2\pi i} \displaystyle\int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} \dfrac{(n/m)^{-s}}{s(1-s)}\, ds \\ \qquad = \dfrac{1}{2\pi i} \displaystyle\int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} \dfrac{\alpha_N(s)\, \bar{\alpha}_M(1-s)}{s(1-s)}\, ds \end{cases}$$

where α_N denotes a partial sum. Consider the integral of the same integrand taken around the rectangle with vertices $\frac{1}{2} \pm iT$, $c \pm iT$. The integrand is analytic within the contour, and the integrals along the two lines $\frac{1}{2} \leqq \sigma \leqq c$ both tend to zero as $T \to \infty$; thus (4.4) is valid with $\frac{1}{2}$ replaced by c. Then as $N \to \infty$, $S_{N,M}$ tends by dominated convergence to

$$S_M = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{\alpha(s)\,\bar{\alpha}_M(1-s)}{s(1-s)}\,ds.$$

Now the conditions (ii) and (iii) on $\alpha(s)$ permit the contour to be moved back to $\sigma = 1-c$, in the same way; in fact,

$$\left| \int_{1-c+iT}^{c+iT} \right| = O\left(|T|^{-2}\right) \sum_1^M |a_n|\, n^{c-1} \int_{1-c}^{c} |\alpha(\sigma + iT)|\, d\sigma = o(1)$$

as $T \to \pm\infty$, by (iii). The integral along $\sigma = 1-c$ so obtained converges absolutely:

$$\left| \int_{1-c-iT}^{1-c+iT} \right| \leqq \int_{-\infty}^{\infty} O\left(\frac{1}{1+|t|^2}\right) O\left(|t|^{1-\varepsilon}\right) dt \cdot \sum_1^\infty |a_n|\, n^{-c} < \infty.$$

Hence by dominated convergence again, S_M tends to a limit as $M \to \infty$. The result follows. The theorem gives no additional information if $a_n \geqq 0$ for all n, for then $\alpha(s)$ has a singularity at $s = \bar{\sigma}$, and (ii) cannot hold.

We summarize this and the last section in

THEOREM 2. *Given a kernel K for which* (2.6) *holds, there is a one-to-one correspondence between w-sequences $\{a_n\}$ and sequence-functions P of $\mathscr{H}_{-1}$ which are expandible as K-series $\sum a_n K_n$.*

5. The K-series of an arbitrary sequence-function

It is easily seen from examples that the K_n do not in general form a complete set in $\mathscr{H}_{-1}$, so that an arbitrary sequence-function F is not necessarily expandible in a K-series. Suppose that (2.6) holds, and that the K_n generate the linear subspace $\mathscr{P}$, with orthogonal complement $\mathscr{Q}$ in $\mathscr{H}_{-1}$. For F we can write $F = P + Q$, $P \in \mathscr{P}$, $Q \in \mathscr{Q}$, so that $P = \sum a_n K_n$ for a sequence $\{a_n\}$ determined by

$$\sigma_n = (P, K_n)_{-1} = (F, K_n)_{-1} \quad (n \geqq 1), \qquad \sigma_0 = 0.$$

We can call $\{a_n\}$ the sequence of "K-coefficients" of F if the a_n satisfy (3.2), i.e. if

$$(F, K_m)_{-1} = \sum_{n=1}^{\infty} a_n [\max(n, m)]^{-1},$$

and $\sum a_n K_n$ the "K-series" of F, writing $F \sim \sum a_n K_n$. The work of §3 shows that the K-coefficients and K-series are uniquely determined. In place of (3.4) we have $\|F\|^2_{-1} \geqq \sum a_n \bar{\sigma}_n$. Clearly, the K-coefficients of an arbitrary sequence-function form a w-sequence. This process fails if (2.6) does not hold.

We now give a characterization of sequence-functions of $\mathscr{P}$ in terms of Fourier series. Let $\{a_n\}$ be an arbitrary real w-sequence, K a real kernel, and let P_N (defined by (3.1)) converge in $\mathscr{H}_{-1}$ to P. The Mellin transform of $x^{-1} P_N^{(-1)}(x)$ is

$$\mathfrak{P}_{N,-1}(s) = \sum_1^N a_n n^{-s} (1-s)^{-1} \mathfrak{K}(s),$$

and $\mathfrak{P}_{N,-1}(\frac{1}{2}+it)$ converges in $L^2(-\infty, \infty)$ to the Mellin transform $\mathfrak{P}_{-1}(s)$ of $x^{-1} P^{(-1)}(x)$[4]:

$$\mathfrak{P}_{-1}(s) = \operatorname*{l.i.m.}_{N\to\infty} (1-s)^{-1} \mathfrak{K}(s) \cdot \sum_1^N a_n n^{-s}.$$

Now the Mellin transform of

$$(nx)^{-1} \sin nx \quad \text{is} \quad -\Gamma(s-1) \cos \tfrac{1}{2} s\pi \cdot n^{-s},$$

and so

$$(1-s)^{-1} \Gamma(s) \cos \tfrac{1}{2} s\pi \cdot \sum_1^N a_n n^{-s} = \int_0^\infty \left\{ \sum_1^N a_n (nx)^{-1} \cos^{(-1)}(nx) \right\} x^{s-1} dx.$$

Then, since $|\Gamma(s) \cos \frac{1}{2} s\pi|$ is constant on $\sigma = \frac{1}{2}$, we have

$$(5.1)\quad \mathfrak{P}_{-1}(s)\, \mathfrak{K}^*(1-s)\, \Gamma(s) \cos \tfrac{1}{2} s\pi = \operatorname*{l.i.m.}_{N\to\infty} \int_0^\infty \left\{ \sum_1^N a_n (nx)^{-1} \cos^{(-1)}(nx) \right\} x^{s-1} dx.$$

But since $\{a_n\}$ is a w-sequence, the sum under the integral sign converges in L^2; and it also converges pointwise. The righthand side may therefore be written as the L^2 Mellin transform of the infinite series.

The argument is reversible. We have

THEOREM 3. *A necessary and sufficient condition that a sequence-function F of $\mathscr{H}_{-1}$ be expandible in the form*

$$F(x) = \lim_{N\to\infty (-1)} \sum_1^N a_n K(nx)$$

for a given kernel K of $\mathscr{H}_{-1}$ is that

$$\mathfrak{F}_{-1}(s)\, \mathfrak{K}^*(1-s)\, \Gamma(s) \cos \tfrac{1}{2} s\pi$$

be the L^2 Mellin transform of the integrated Fourier series

$$\sum_1^\infty a_n (nx)^{-1} \cos^{(-1)}(nx),$$

where $\{a_n\}$ is a w-sequence.

---

[4]) We use (2.2); hence the necessity that P_N be real.

A similar relation connects $\mathfrak{F}_{-1}(s)\,\mathfrak{K}^*(1-s)\,\Gamma(s)\sin\frac{1}{2}s\pi$ with a sine series. The Mellin transform relation in either case is equivalent to (1.2), interpreted as

$$\mathfrak{F}_{-1}(s) = (1-s)^{-1}\,\mathfrak{F}(s) = \underset{N\to\infty}{\text{l.i.m.}}\,(1-s)^{-1}\,\mathfrak{K}(s)\sum_{n=1}^{N} a_n n^{-s}.$$

In this sense, (1.2) gives an alternate characterization of the space $\mathscr{P}$ of sequence-functions developable in series form.

6. The K-transform of a K-series

Just as in Schwartz' spaces of distributions the Fourier transform of a Fourier series is a sum of delta distributions[5]), so here the K-transform of a K-series is a sum of delta sequence-functions.

To prove this, we consider first the case when the kernel is a real principal sequence-function k (i.e. $k\in\mathscr{G}_{-1}$), and $k^*=k$. It is known ([5], §10) that

$$\delta_n(x,\xi) = \int_0^n k(xt)\,k(\xi t)\,dt \qquad (\xi>0)$$

converges in $\mathscr{H}_{-1}$ to a delta sequence-function Δ_ξ if (i) $\int_x^\infty t^{-1}k(t)\,dt = o(x^{-\frac{1}{2}})$ as $x\to\infty$, and (ii) $k(x)\in L^2(0,X)$, $k(x)\,x^{-\frac{1}{2}}\in L(0,X)$ for all positive X.

We obtain a sequence (φ_n), $\varphi_n\in L^2$, for k by taking

$$\varphi_n(x) = k(x) \qquad (0<x<n),\quad 0 \quad (x>n).$$

For then $\varphi_n\in L^2$ for all n, by (ii), and its k-transform is

$$\underset{X\to\infty}{\text{l.i.m.}}\int_0^X \varphi_n(t)\,k(xt)\,dt = \delta_n(x,1).$$

Also, since $\|\delta_n(x,1)-\Delta_1(x)\|_{-1}\to 0$ as $n\to\infty$,

$$\|\varphi_n-\varphi_m\|_{-1} = \|\delta_n(x,1)-\delta_m(x,1)\|_{-1}\to 0 \qquad (n,m\to\infty)\,^6),$$

and so (φ_n) defines a sequence-function in $\mathscr{H}_{-1}$. Now if $m<n$,

$$\|\varphi_n-\varphi_m\|_{-1}^2 = \int_m^n x^{-2}\left|\int_m^x k(t)\,dt\right|^2 dx + \int_n^\infty x^{-2}\left|\int_m^n k(t)\,dt\right|^2 dx.$$

Each term on the righthand side must approach zero; thus

$$\|k-\varphi_m\|_{-1}^2 = \int_m^\infty x^{-2}\left|\int_m^x k(t)\,dt\right|^2 dx \to 0 \qquad (m\to\infty),$$

and hence (φ_m) determines the sequence-function k.

Let T be the transformation

$$T\{f,x\} = \underset{X\to\infty}{\text{l.i.m.}}\int_0^X f(t)\,k(xt)\,dt,$$

---

[5]) Cf. [4], p. 58.

[6]) By the Parseval relation $\|f\|_{-1} = \|Tf\|_{-1}$ for a kernel related to k (T as defined below); see [5], (11.5).

mapping L^2 onto itself, and let T_{-1} be its extension to $\mathscr{H}_{-1}$, defined by continuity ([5], §11). Then

$$T_{-1}\{k(mt), x\} = \lim_{n\to\infty(-1)} T\{\varphi_n(mt), x\}$$
$$= \lim_{n\to\infty(-1)} \int_0^{n/m} k(mt)\, k(xt)\, dt = \Delta_m(x).$$

Let $p_N = \sum_1^N a_m k(mx)$ converge in $\mathscr{H}_{-1}$ to P. Then

$$T_{-1}(P) = \lim_{N\to\infty(-1)} T_{-1}(p_N) = \lim_{N\to\infty(-1)} \sum_1^N a_m \Delta_m.$$

We have

THEOREM 4. *If k is a real kernel, for which $k^* = k$, and satisfying* (i) *and* (ii) *above, the k-transform of*

$$\sum_1^\infty a_m k(mx) \quad \textit{is} \quad \sum_1^\infty a_m \Delta_m(x).$$

To obtain a similar result for an arbitrary real kernel K we must work differently, since the previous sequence for Δ_ξ is no longer defined. We take a new sequence (ψ_n) for K, defining ψ_n through its Mellin transform

$$\Psi_n(s) = \mathfrak{K}(s) \quad (|t| < n), \quad 0 \quad (|t| > n) \quad (s = \tfrac{1}{2} + it).$$

Then ψ_n is real and belongs to L^2 for each n; (2.2) shows that $\|\psi_n - K\|_{-1} \to 0$. If χ_n is the K^*-transform of ψ_n, its Mellin transform is, by (2.5),

$$\Psi_n(1-s)\, \mathfrak{K}^*(s) = 1 \quad (|t| < n), \quad 0 \quad (|t| > n).$$

Now the Mellin transform of $x^{-1}\Delta_m^{(-1)}(x) = x^{-1}(x > m)$, 0 $(0 < x < m)$ is $m^{s-1}(1-s)^{-1}$.[7] Thus by (2.2) again,

$$\left\| \frac{1}{m}\chi_n\left(\frac{x}{m}\right) - \Delta_m(x) \right\|_{-1}^2 = \frac{1}{2\pi}\left\{ \int_{-\infty}^{-n} + \int_n^\infty \right\} |(1-s)^{-1} m^{s-1}|^2\, dt \to 0.$$

Then

$$T_{-1}^*\{K_m, x\} = \lim_{n\to\infty(-1)} T^*\{\psi_n(mt), x\}$$
$$= \lim_{n\to\infty(-1)} \frac{1}{m}\chi_n\left(\frac{x}{m}\right) = \Delta_m(x),$$

and we find

$$T_{-1}^*\left\{\sum_1^\infty a_m K_m\right\} = \sum_1^\infty a_m \Delta_m.$$

7. Examples

1°. *Fourier series.* Take $k(x) = 2\cos 2\pi x$; the series $\sum a_n 2\cos 2\pi n x$, if it converges pointwise, defines a function $p(x)$ say, both $p(x)$ and $p^{(-1)}(x)$ having period 1, and the relation

$$p^{(-1)}(x) = p^{(-1)}(1-x) \qquad (0 \leqq x \leqq 1) \tag{7.1}$$

[7] Cf. [5], Theorem 5.

holding. Conversely, if f is a function having the same properties, then a_m as given by (3.3) is equal to the usual Fourier coefficient: we obtain the identity (which can be proved independently)

$$a_m = \frac{4m}{\pi} \int_0^\infty x^{-2} f^{(-1)}(x) \sin^2 \pi x \sin 2\pi m x \, dx = \int_0^1 f(x) \cos 2\pi m x \, dx.$$

Similarly, for periodic functions satisfying $f^{(-1)}(x) = -f^{(-1)}(1-x)$ for $0 \leqq x \leqq 1$, instead of (7.1), we find

$$-\frac{4m}{\pi} \int_0^\infty x^{-2} f^{(-1)}(x) \sin^2 \pi x \cos 2\pi m x \, dx = \int_0^1 f(x) \sin 2\pi m x \, dx.$$

2^0. A function f equal to the sum of its k-series is, by (1.2), closely associated with the Dirichlet series $\alpha(s) = \sum a_n n^{-s}$; the form of the more obvious examples is dictated by this series.

Consider the kernel

$$k(x) = 4K_0(4\pi x^{\frac{1}{2}}) - 2\pi Y_0(4\pi x^{\frac{1}{2}}), \quad \text{for which} \quad \mathfrak{K}(s) = \frac{\zeta^2(1-s)}{\zeta^2(s)}.$$

Now by using the functional equation of the Riemann zeta function $\zeta(s)$, we can show that

$$(2\pi)^4 \frac{\zeta^2(-s-1)}{s^2(1+s)^2} = \frac{\zeta^2(1-s)}{\zeta^2(s)} \cdot \zeta^2(s+2).$$

This has the form (1.2), with[8])

$$\alpha(s) = \zeta^2(s+2) = \sum_1^\infty \frac{d(n)}{n^2} n^{-s} \qquad (\mathfrak{R}(s) > -1);$$

here $d(n)$ is the number of divisors of n, $\{d(n)\, n^{-2}\}$ is a w-sequence, and therefore $F \in \mathscr{H}_{-1}$. Thus

$$x^{-1} F^{(-1)}(x) = \frac{1}{2\pi i} \int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} \frac{\mathfrak{F}(s)}{1-s} x^{-s} \, ds$$

$$= (2\pi)^4 \cdot \frac{1}{2\pi i} \int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} \frac{\zeta^2(-s-1)\, x^{-s}}{s^2(1+s)^2(1-s)} \, ds.$$

To evaluate this integral we first put $s = 1-z$, and then move the contour to the right; this can be justified by using known uniform order relations for $\zeta(s)$[9]). We obtain

$$\frac{x^{-1} F^{(-1)}(x)}{(2\pi)^4} = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{\zeta^2(z-2)\, x^{z-1}}{(1-z)^2(2-z)^2 z} \, dz - R,$$

where c is taken so large that $\sum d(n)\, n^{2-c}$ converges, i.e. $c > 3$, and R is the sum of the residues of the integrand at $z = 1, 2$ and 3. Term-by-term

[8]) For properties of the arithmetic functions $d(n)$ and $r(n)$ see [2], Ch. XVII.

[9]) Cf. [8], Ch. I.

integration of the series obtained by substituting for $\zeta^2(z-2)$ gives for the integral

$$(4x)^{-1}\sum_{n\leq x} d(n)\{n^2+4nx-5x^2+2x(2n+x)\log(x/n)\},$$

and

$$R=\frac{1}{144}(1+18x+12x^2)\log x+\frac{1-24\zeta'(-1)}{144}+$$
$$+\frac{x}{16}(4\log 2\pi-5)+\frac{x^2}{18}(3\gamma-5).$$

Thus $F^{(-1)}(x)$ is absolutely continuous, and $F^{(-1)}(0+)=0$; by differentiation we obtain the relation

$$(7.2)\quad \begin{cases} -\sum_{1\leq n\leq x} d(n)\{2(x-n)+(x+n)\log(n/x)\}-\frac{1}{4}\left(\frac{1}{36}+x+x^2\right)\log x+ \\ \quad +\frac{12\zeta'(-1)-1}{72}+\frac{x}{2}(1-\log 2\pi)+\frac{x^2}{4}(3-2\gamma) \\ =(2\pi)^{-4}\sum_{m=1}^{\infty}\frac{d(m)}{m^2}\left[4K_0\left(4\pi(mx)^{\frac{1}{2}}\right)-2\pi Y_0\left(4\pi(mx)^{\frac{1}{2}}\right)\right]. \end{cases}$$

3⁰. The Dirichlet series $L(s)=1^{-s}-3^{-s}+5^{-s}-\cdots$ satisfies the functional equation

$$L(s)=\pi^{s-\frac{1}{2}}2^{1-2s}\frac{\Gamma(1-\frac{1}{2}s)}{\Gamma(\frac{1}{2}+\frac{1}{2}s)}L(1-s);$$

hence if

$$\alpha(s)=4\zeta(s+1)L(s+1)=\sum_1^{\infty}\left(r(n)/n\right)n^{-s}\qquad\left(\Re(s)>0\right),$$

where $r(n)$ is the number of representations of n as a sum of two squares, we can deduce

$$-\pi^2\frac{\alpha(-1-s)}{s^2}=\pi^{1-2s}\frac{\Gamma(s)}{\Gamma(1-s)}\cdot\alpha(s).$$

From this identity as a form of (1.2) we derive the relation

$$(7.3)\quad \begin{cases} \sum_{1\leq n\leq x} r(n)\log(n/x)-\log x-\log 2\pi-2L'(0)+\pi x \\ \qquad =\pi^{-1}\sum_{m=1}^{\infty} r(m)\,m^{-1}J_0\left(2\pi(mx)^{\frac{1}{2}}\right). \end{cases}$$

In examples 2⁰ and 3⁰, the series converge in the first instance in the norm of $\mathscr{H}_{-1}$; but they also converge pointwise, and (7.2) and (7.3) hold for all positive x. In each case, (4.2) holds for the sequence $\{a_n\}$. A formal differentiation of (7.3) gives Voronöi's identity.

4⁰. There are other examples in the same vein as 2⁰ and 3⁰. The method was used by APOSTOL [10] to prove Landau's identity,

$$(7.4)\quad \begin{cases} \frac{1}{q!}\sum_{0\leq n\leq x} a(n)(x-n)^q=\frac{\varrho\,\Gamma(\varkappa)}{\Gamma(\varkappa+q+1)}x^{\varkappa+q}+ \\ \qquad +\gamma\left(\frac{\lambda}{2\pi}\right)^q x^{\frac{1}{2}(\varkappa+q)}\sum_{n=1}^{\infty}\frac{a(n)}{n^{\frac{1}{2}(\varkappa+q)}}J_{\varkappa+q}\left(\frac{4\pi}{\lambda}(nx)^{\frac{1}{2}}\right); \end{cases}$$

[10] [1]; $\gamma=\pm 1$, $a(0)=-\varphi(0)$, q is a positive integer.

where $\sum a(n)n^{-s}$ is assumed to converge absolutely for $\sigma>\varkappa$, defining a function $\varphi(s)$, analytic except possibly for a simple pole at $s=\varkappa$ with residue ϱ, which satisfies

$$\left(\frac{2\pi}{\lambda}\right)^{-s}\Gamma(s)\,\varphi(s)=\gamma\left(\frac{2\pi}{\lambda}\right)^{s-\varkappa}\Gamma(\varkappa-s)\,\varphi(\varkappa-s). \tag{7.5}$$

Here

$$a_n=a(n)\,n^{-\frac{1}{2}(\varkappa+q)},\quad k(x)=k^*(x)=\frac{2\pi}{\lambda}J_{\varkappa+q}\left(\frac{4\pi}{\lambda}x^{\frac{1}{2}}\right),$$

and (7.5) shows that (1.2) has the form

$$\gamma\left(\frac{2\pi}{\lambda}\right)^{q+1}\frac{\alpha(-q-s)\,\Gamma(\frac{1}{2}\varkappa-\frac{1}{2}q-s)}{\Gamma(\frac{1}{2}\varkappa+\frac{1}{2}q+1-s)}=\left(\frac{2\pi}{\lambda}\right)^{1-2s}\frac{\Gamma(s+\frac{1}{2}\varkappa+\frac{1}{2}q)}{\Gamma(1-s+\frac{1}{2}\varkappa+\frac{1}{2}q)}\cdot\alpha(s).$$

The series in (7.4) converges absolutely if $q>\varkappa-\frac{1}{2}$; while (4.2) holds (so that $\{a_n\}$ is a w-sequence) if $q>\varkappa-1$. For a sequence satisfying the conditions of Theorem 1 we may have convergence in $\mathscr{H}_{-1}$ for smaller q.

5⁰. Let $k(x)=x^{\frac{1}{2}}J_\nu(x)$ and take $a_n=n^{-\nu-\frac{1}{2}}$. Then $\alpha(s)=\zeta(s+\nu+\frac{1}{2})$, and the functional equation of the Riemann zeta function shows that (1.2) can be written

$$2^{s-\frac{1}{2}}\pi^{s+\nu}\frac{\Gamma(\frac{1}{4}-\frac{1}{2}\nu-\frac{1}{2}s)}{\Gamma(\frac{3}{4}+\frac{1}{2}\nu-\frac{1}{2}s)}\zeta(\tfrac{1}{2}-s-\nu)=2^{s-\frac{1}{2}}\frac{\Gamma(\frac{1}{4}+\frac{1}{2}\nu+\frac{1}{2}s)}{\Gamma(\frac{3}{4}+\frac{1}{2}\nu-\frac{1}{2}s)}\cdot\zeta(s+\nu+\tfrac{1}{2}),$$

By the method of 2⁰, we obtain the Schlömilch series

$$\sum_{n=1}^{\infty}n^{-\nu-\frac{1}{2}}(nx)^{\frac{1}{2}}J_\nu(nx)=\frac{2^{1-\nu}\pi^{\frac{1}{2}}}{\Gamma(\nu+\frac{1}{2})}x^{\frac{1}{2}-\nu}\sum_{0<2\pi n\le x}(x^2-4\pi^2n^2)^{\nu-\frac{1}{2}}+$$

$$+\frac{2^{-\nu}\pi^{\frac{1}{2}}x^{\nu-\frac{1}{2}}}{\Gamma(\nu+\frac{1}{2})}-\frac{x^{\nu+\frac{1}{2}}}{2^{\nu+1}\Gamma(\nu+1)}.$$

This identity is also a case of Poisson's summation formula [11]). The series converges absolutely if $\nu>\frac{1}{2}$, and in $\mathscr{H}_{-1}$ if $\nu>0$, since $\{a_n\}$ is then a w-sequence, by (4.2).

6⁰. The functional equation of $L(s)$ (3⁰) can be written

$$L(1-s)=\mathfrak{S}(s)\,L(s), \tag{7.6}$$

where $\mathfrak{S}(s)$ is the Mellin transform of the kernel $\sin\frac{1}{2}\pi x$. With $\alpha(s)=L(s)$, $\{a_n\}$ is the w-sequence $1, 0, -1, 0, \ldots$ ((4.2) does not hold). Hence $F\in\mathscr{H}_{-1}$ and by the previous method we find

$$x^{-1}F^{(-1)}(x)=x^{-1}\sum_{1\le 2n+1<x}(-1)^n.$$

But (cf. §6), $x^{-1}\Delta_m^{(-1)}(x)=0$ $(0<x<m)$, x^{-1} $(x>m)$. Thus

$$F(x)=\sum_{n=0}^{\infty}(-1)^n\Delta_{2n+1}(x),$$

---

[11]) Cf. [7], p. 65.

and we have

$$\sum_{n=0}^{\infty}(-1)^n \Delta_{2n+1}(x) = \sum_{n=0}^{\infty}(-1)^n \sin(n+\tfrac{1}{2})\pi x.$$

Theorem 4 shows that this sequence-function is self-reciprocal with respect to the transformation with kernel $\sin \frac{1}{2}\pi x$, as (7.6) implies.

References

[1] Apostol, T. M.: Identities involving the coefficients of certain Dirichlet series. Duke Math. J. **18**, 517–525 (1951).

[2] Bateman Manuscript Project: Higher transcendental functions, Vols 2, 3. McGraw-Hill 1953, 1955.

[3] — Tables of integral transforms, Vol. 1. McGraw-Hill 1954.

[4] Lighthill, M. J.: Fourier analysis and generalized functions. Cambridge 1958.

[5] Miller, J. B.: Hilbert spaces of generalized functions extending $L^2(I)$. J. Aust. Math. Soc. **1**, 281—298 (1960).

[6] Oppenheim, A.: Some identities in the theory of numbers. Proc. London Math. Soc. **26**, 295–350 (1927).

[7] Titchmarsh, E. C.: Theory of Fourier integrals, 2nd ed. Oxford 1948.

[8] — The zeta-function of Riemann. Cambridge Tract No. 26. Cambridge 1930.

Australian National University, Canberra, A.C.T.

(Eingegangen am 28. Juni 1960)

MACDONALD, I. D.
Math. Zeitschr. 76, 270—282 (1961)

On Certain Varieties of Groups

By
I. D. MACDONALD

A variety is a class of groups defined by laws, or identical relations. Thus the abelian groups form the variety associated with the law

$$(x, y) = 1,$$

which is to be understood as holding good for any pair x, y of elements in each group of the variety.

The varieties discussed here are those associated with the laws

(a) $$(x, y_1, \ldots, y_n, x) = 1,$$

(b) $$(x, y_1, \ldots, y_n; x, z_1, \ldots, z_m) = 1$$

respectively, where m and n are fixed positive integers. We prove that the groups G of the variety (a) are nilpotent of class $n+2$ or less and that the term G_{n+1} of the lower central series has exponent 3 — a well-known result when $n=1$. For any group G of the variety (b) with $n \geqq m$, we are able to show that the term G_m of the lower central chain is nilpotent of class $n-m+2$ or less; therefore G is soluble. Additional results are proved in the cases $m=n=1$ and $m+1=n=2$.

Corresponding theorems would hold for varieties of Lie rings, the most interesting being the analogue of the lemma in §2. We do not state any of these results explicitly.

1. Notation and basic facts

The simple commutator $(a_1, a_2, \ldots, a_n)$ in the group G is defined in the standard way:

$$(a_1, a_2) = a_1^{-1} a_2^{-1} a_1 a_2,$$

and for all $n > 2$

$$(a_1, a_2, \ldots, a_n) = \big((a_1, a_2, \ldots, a_{n-1}), a_n\big).$$

Further we shall put

$$(a_1, a_2; b_1, b_2) = \big(a_1, a_2, (b_1, b_2)\big).$$

The symbol 1 will denote the unit element or subgroup of whatever group G is under consideration, and for any elements a and b of G we use a^b for $b^{-1} a b$. Thus $(a, b) = a^{-1} a^b$.

The following list of commutator identities is also standard:

$$(1.1) \qquad (a, b)^{-1} = (b, a);$$

$$(1.2) \qquad (a, b) = (b, a^{-1})^a = (b^{-1}, a)^b;$$

$$(1.3) \qquad \begin{cases} (a\,b, c) = (a, c)^b\,(b, c), \\ (a, b\,c) = (a, c)\,(a, b)^c; \end{cases}$$

$$(1.4) \qquad (a\,b, c\,d) = (a, d)^b\,(b, d)\,(a, c)^{b\,d}\,(b, c)^d;$$

$$(1.5) \qquad (c, a^{-1}, b)^a\,(a, b^{-1}, c)^b\,(b, c^{-1}, a)^c = 1;$$

$$(1.6) \qquad (c, a, b)\,(a, b, c)\,(b, c, a) = (a, c)\,(b, c)\,(b, a)^c\,(c, a)\,(c, b)^a\,(a, b)^c\,(c, b)\,(b, c)^a.$$

Identity (1.3) can be extended as follows:

$$(1.7) \qquad (a\,b, c_1, \ldots, c_n) = \big(\ldots (a, c_1)^b, c_2\big)^{(b, c_1)}, \ldots, c_n\big)^{(b, c_1, \ldots, c_{n-1})}(b, c_1, \ldots, c_n).$$

This form is easily established by induction. We shall also use another simple consequence of (1.3):

$$(1.8) \qquad (a, b)^c = (a, c)^{-1}\,(a, b\,c).$$

If H and K are subgroups of G then the subgroup generated by the commutators (h, k) for all $h \in H$ and all $k \in K$ is denoted by (H, K). The terms G_i of the lower central chain of G are defined by

$$G_0 = G, \qquad G_{n+1} = (G_n, G)$$

for all $n \geqq 0$. The commutator subgroup G_1 of G will usually be written as G'. The subgroup G'' is defined to be (G', G'). The centre of G will be called Z or, more explicitly, $Z(G)$.

2. The variety $(x, y_1, \ldots, y_n, x) = 1$

THEOREM 1. *Every group G of the variety with the law*

$$(2.1) \qquad (x, y_1, \ldots, y_n, x) = 1$$

where n is a fixed positive integer, has the properties

(i) *G is nilpotent of class $n+2$ or less; and*

(ii) *G_{n+1} has exponent three*[1]).

We note that this theorem is known for $n = 1$, being due to F. W. LEVI [2]. The proof is also to be found in the book [1] of M. HALL, pp. 321–2; and elsewhere. Further, for $n = 1$ examples of groups are known in which the class is just 3. We leave open the question whether the bound $n+2$ in (i) and the exponent three in (ii) are best possible.

LEVI's result is to be used as the basis of an inductive proof of Theorem 1. In this the following lemma will be needed.

[1]) Since first submitting this paper I have noticed results of H. MEIER-WUNDERLI in „Über die Gruppen mit der identischen Relation $(x_1, x_2, \ldots, x_n) = (x_2, x_3, \ldots, x_n, x_1)$ $(n \geqq 3)$", Vjschr. naturf. Ges. Zürich **94**, 211–218 (1949). Theorem 1 above seems to be a stronger version of them.

LEMMA. *In any group G the commutator $(a_1, a_2, \ldots, a_n, b)$ can be expressed as the product of 4^{n-1} commutators of the form $(b, a_1^*, a_2^*, \ldots, a_n^*)^{\pm 1}$ where each a_i^* is a conjugate of some $a_j^{\pm 1}$.*

PROOF[2]). We proceed by induction on n; the result for $n=1$ is given at once by (1.1). Suppose that the result is true for a fixed positive integer N, then we shall prove it in the case of the commutator $(a_1, \ldots, a_{N+1}, b)$.

Let us put $(a_1, \ldots, a_N) = c$. Identity (1.5) gives us

$$(c, a_{N+1}, b) = (b, c^{-1}, a_{N+1}^{-1})^{-c a_{N+1}} (a_{N+1}^{-1}, b^{-1}, c)^{-b a_{N+1}}. \tag{2.2}$$

But by (1.2)

$$(b, c^{-1}, a_{N+1}^{-1})^{-c a_{N+1}} = (c, b, a_{N+1}^{-c})^{-a_{N+1}}, \tag{2.3}$$

and by the inductive hypothesis

$$(c, b) = \prod_{i=1}^{2^{2N-2}} (b, a_{i1}, \ldots, a_{iN})^{\pm 1},$$

the a_{ij} being certain conjugates of the $a_k^{\pm 1}$. Repeated application of (1.3), (1.2), (1.1) and (1.8) will give

$$(c, b, a_{N+1}^{-c})^{-a_{N+1}} = \prod_{i=1}^{2^{2N-2}} (b, a'_{i1}, \ldots, a'_{i,N+1})^{\pm 1}, \tag{2.4}$$

where the a'_{ij} are certain conjugates of the $a_k^{\pm 1}$. We also have

$$\begin{cases} (a_{N+1}^{-1}, b^{-1}, c)^{-b a_{N+1}} = \big((b, a_{N+1}^{-1}), c^b\big)^{-a_{N+1}} \\ \qquad\qquad\qquad\qquad = \big(c^b, (b, a_{N+1}^{-1})\big)^{a_{N+1}} \end{cases} \tag{2.5}$$

by (1.2) and (1.1). The inductive hypothesis and (1.8) give

$$\begin{cases} \big(c^b, (b, a_{N+1}^{-1})\big)^{a_{N+1}} = \prod_{i=1}^{2^{2N-2}} \big((b, a_{N+1}^{-1}), a''_{i1}, \ldots, a''_{iN}\big)^{\pm a_{N+1}} \\ \qquad\qquad\qquad\qquad = \prod_{i=1}^{2^{2N-2}} (b, a_{i1}^*, \ldots, a_{i,N+1}^*)^{\pm 1}, \end{cases} \tag{2.6}$$

where the a''_{ij} and the a_{ij}^* are certain conjugates of the $a_k^{\pm 1}$.

Now (2.2), (2.3), (2.4), (2.5) and (2.6) show that $(a_1, \ldots, a_{N+1}, b)$ is the product of 4^N commutators of the form $(b, a_1^*, \ldots, a_{N+1}^*)^{\pm 1}$ where the a_i^* are certain conjugates of the $a_j^{\pm 1}$.

This completes the inductive proof of the lemma.

PROOF OF THEOREM 1. First we verify that for all n (2.1) is equivalent to

$$\big((x, y_1, \ldots, y_n)^x, x\big) = 1. \tag{2.7}$$

For by (1.8)

$$\big((x, y_1, \ldots, y_{n-1}), y_n\big)^x = \big((x, y_1, \ldots, y_{n-1}), z\big)^{-1} \big((x, y_1, \ldots, y_{n-1}), y_n z\big),$$

and application of (2.1) gives (2.7) immediately.

[2]) Since first submitting this paper I have noticed a result of C. R. B. WRIGHT which contains the central idea of this proof. See "On groups of exponent four with generators of order two", Pacific Journal Math. **10**, 1097–1105 (1960), Lemma 3.2.

We now assume that when n is a fixed positive integer N the groups of the variety (2.1) have class $N+2$ or less; and consider a group G of the variety

$$(x, y_1, \ldots, y_{N+1}, x) = 1. \tag{2.8}$$

By (2.7) we have the law

$$\left((x, y_1, \ldots, y_{N+1})^x, x\right) = 1. \tag{2.9}$$

We put $x = uv$ in (2.8) and expand the resulting law

$$(uv, y_1, \ldots, y_{N+1}, uv) = 1. \tag{2.10}$$

By (1.7) we have

$$(uv, y_1, \ldots, y_{N+1}) = \left(\left(\ldots (u, y_1)^v, \ldots, y_{N+1}\right)^{(v, y_1, \ldots, y_N)} (v, y_1, \ldots, y_{N+1})\right).$$

Therefore

$$(uv, y_1, \ldots, y_{N+1}, uv) = c_1 c_2 c_3 c_4 \tag{2.11}$$

by (1.4), where

$$\begin{aligned}
c_1 &= \left(\left(\ldots (u, y_1)^v, \ldots, y_{N+1}\right)^{(v, y_1, \ldots, y_N)}, v\right)^{(v, y_1, \ldots, y_{N+1})}, \\
c_2 &= (v, y_1, \ldots, y_{N+1}, v), \\
c_3 &= \left(\left(\ldots (u, y_1)^v, \ldots, y_{N+1}\right)^{(v, y_1, \ldots, y_N)}, u\right)^{(v, y_1, \ldots, y_{N+1})v}, \\
c_4 &= (v, y_1, \ldots, y_{N+1}, u)^v.
\end{aligned}$$

We can simplify the c_i with the aid of (2.9):

$$\begin{aligned}
c_2 &= c_3 = 1, \\
c_1 &= \left(\left(\ldots (u, y_1)^v, \ldots, y_{N+1}\right)^{(v, y_1, \ldots, y_N)}, v\right), \\
c_4 &= (v, y_1, \ldots, y_{N+1}, u).
\end{aligned}$$

Now we rewrite (2.10) and (2.11) as[3])

$$(v, y_1, \ldots, y_{N+1}, u) = (u', y_1', \ldots, y_{N+1}', v)^{-1}, \tag{2.12}$$

where $u', y_1', \ldots, y_{N+1}'$ are certain conjugates of $u, y_1, \ldots, y_{N+1}$ respectively.

Put $y_{N+1} = v$ in (2.12). Because of the lemma $(u', y_1', \ldots, y_{N+1}')$ can be expressed as the product of conjugates of elements like $(v, w_1, \ldots, w_{N+1})^{\pm 1}$.

[3]) We could in fact simplify (2.12) by showing that $c_1 = (u, y_1, \ldots, y_{N+1}, v)$ and a similar result is proved after (3.5) below.

Hence, in Meier-Wunderli's variety the relations

$$\begin{aligned}
(x_1, x_2, \ldots, x_n) &= (x_2, x_3, \ldots, x_n, x_1), \\
(x_1, x_2, \ldots, x_n) &= (x_n, x_2, \ldots, x_{n-1}, x_1)^{-1},
\end{aligned}$$

both hold. From these and the fact that

$$(x_1, x_2, \ldots, x_n)^3 = 1,$$

it is possible to show that for even n exceeding 3

$$(x_1, x_2, \ldots, x_n) = 1.$$

Then (2.9) shows that the right-hand side of (2.12) is 1, that is

$$(v, y_1, \ldots, y_N, v, u) = 1. \tag{2.13}$$

A group G with the law (2.13) has the property that $(v, y_1, \ldots, y_N, v)$ lies in the centre Z. In other words G/Z belongs to the variety (2.1), with $n = N$. As G/Z is nilpotent of class $N+2$ or less by the inductive hypothesis, G is nilpotent of class $N+3$ or less. Moreover by the inductive hypothesis G/Z satisfies the law

$$(z_1, \ldots, z_{N+2})^3 = 1$$

and so G satisfies

$$((z_1, \ldots, z_{N+2})^3, z) = 1. \tag{2.14}$$

On applying (1.3) and (2.9) to this we find that

$$(z_1, \ldots, z_{N+2}, z)^3 = 1. \tag{2.15}$$

Thus G_{N+2} has exponent 3, as it is central in G.

This completes the inductive proof of Theorem 1.

3. The variety $(x, y_1, \ldots, y_n; x, z_1, \ldots, z_m) = 1$

To establish the theorem of this section both the method of proof, and the result, of Theorem 1 will be used.

THEOREM 2. *Every group G of the variety with the law*

$$(x, y_1, \ldots, y_n; x, z_1, \ldots, z_m) = 1, \tag{3.1}$$

where $n \geqq m$ and $n > 1$, has G_m nilpotent of class $n - m + 2$ or less.

PROOF. We note that (1.8) shows that (3.1) is equivalent to the law

$$((x, y_1, \ldots, y_n)^s, (x, z_1, \ldots, z_m)^t) = 1. \tag{3.2}$$

In this and later sections we shall use without explicit reference the fact that (3.1) implies

$$(x, y_1, \ldots, y_{n+p}; x, z_1, \ldots, z_m) = 1$$

for all $p \geqq 0$, whether $n \geqq m$ or not. This is a consequence of (3.2) and the fact that $(x, y_1, \ldots, y_{n+p})$ can be written as the product of 2^p conjugates of the commutator $(x, y_1, \ldots, y_n)$ and its inverse.

The relation

$$(uv, y_1, \ldots, y_n; uv, z_1, \ldots, z_m) = 1, \tag{3.3}$$

derived from (3.1), has to be expanded in the way that (2.10) was expanded. For the sake of a later application we do not exclude the case $n = 1$ yet. We note (1.7), then apply (1.4) to the left-hand side of (3.3):

$$(uv, y_1, \ldots, y_n; uv, z_1, \ldots, z_m) = c_1 c_2 c_3 c_4 \tag{3.4}$$

where

$$c_1 = \left((\dots (u, y_1)^v, \dots, y_n)^{(v, y_1, \dots, y_{n-1})}, (v, z_1, \dots, z_m)\right)^g,$$

$$c_2 = (v, y_1, \dots, y_n; v, z_1, \dots, z_m),$$

$$c_3 = \left((\dots (u, y_1)^v, \dots, y_n)^{(v, y_1, \dots, y_{n-1})}, (\dots (u, z_1)^v, \dots, z_m)^{(v, z_1, \dots, z_{m-1})}\right)^{g\,h},$$

$$c_4 = \left((v, y_1, \dots, y_n), (\dots (u, z_1)^v, \dots, z_m)^{(v, z_1, \dots, z_{m-1})}\right)^h,$$

$$g = (v, y_1, \dots, y_n),$$

$$h = (v, z_1, \dots, z_m).$$

With the aid of (3.2) we find easily that

$$c_2 = c_3 = 1,$$

$$c_1 = \left((\dots (u, y_1)^v, \dots, y_n)^{(v, y_1, \dots, y_{n-1})}, (v, z_1, \dots, z_m)\right),$$

$$c_4 = \left((v, y_1, \dots, y_n), (\dots (u, z_1)^v, \dots, z_m)^{(v, z_1, \dots, z_{m-1})}\right);$$

and by (3.3) and (3.4)

$$c_4 = c_1^{-1}. \tag{3.5}$$

The next step consists in showing that

$$c_4 = (v, y_1, \dots, y_n; u, z_1, \dots, z_m).$$

We have

$$c_4 = (v, y_1, \dots, y_n); \left((u, z_1)^{w_1}, z_2^{w_2}, \dots, z_m^{w_m}\right) \tag{3.6}$$

where we have introduced the abbreviations

$$v_1 = v, \qquad v_i = (v, z_1, \dots, z_{i-1})$$

for $2 \leqq i \leqq m$ and

$$w_i = v_i\, v_{i+1} \dots v_m$$

for $1 \leqq i \leqq m$. Suppose that for some i such that $1 \leqq i < m$ we have succeeded in showing that

$$c_4 = \left(v, y_1, \dots, y_n; ((u, z_1, \dots, z_i)^{w_i}, z_{i+1}^{w_{i+1}}, \dots, z_m^{w_m})\right). \tag{3.7}$$

Using the fact that $a^b = (b, a^{-1})\, a$ we find

$$\begin{aligned}(u, z_1, \dots, z_i)^{w_i} &= \left(v_i, (u, z_1, \dots, z_i)^{-1}\right)^{w_{i+1}} (u, z_1, \dots, z_i)^{w_{i+1}} \\ &= (u, z_1, \dots, z_i, v_i)^e\, (u, z_1, \dots, z_i)^{w_{i+1}}\end{aligned}$$

for a certain element e in G by (1.2). Now we apply (1.7):

$$\begin{aligned}&\left((u, z_1, \dots, z_i)^{w_i}, z_{i+1}^{w_{i+1}}, \dots, z_m^{w_m}\right) \\ &\qquad = (u', z_1', \dots, z_i', v_i', z_{i+1}', \dots, z_m')\, \left((u, z_1, \dots, z_i)^{w_{i+1}}, z_{i+1}^{w_{i+1}}, \dots, z_m^{w_m}\right).\end{aligned}$$

where x' denotes a certain conjugate of x. If we can show that $(v, y_1, \ldots, y_n)$ commutes with the first factor on the right-hand side of the last equality then by (3.7) and (1.3) we obtain

$$c_4 = \left(v, y_1, \ldots, y_n; \left((u, z_1, \ldots, z_i)^{w_{i+1}}, z_{i+1}^{w_{i+1}}, \ldots, z_m^{w_m}\right)\right). \tag{3.8}$$

But $(u', z_1', \ldots, z_i', v_i', z_{i+1}', \ldots, z_m')$ can be written as the product of commutators of the forms $(v_i', u^*, z_1^*, \ldots, z_m^*)^{\pm 1}$ by the lemma of §2 and (1.8), and so must commute with $(v, y_1, \ldots, y_n)$ by (3.2).

Compare (3.7) with (3.8). Induction on i at once gives

$$c_4 = \left(v, y_1, \ldots, y_n; (u, z_1, \ldots, z_m)^{w_m}\right); \tag{3.9}$$

but here as before

$$\begin{aligned}(u, z_1, \ldots, z_m)^{w_m} &= \left(v_m, (u, z_1, \ldots, z_m)^{-1}\right)(u, z_1, \ldots, z_m)\\ &= (u, z_1, \ldots, z_m, v_m)^f (u, z_1, \ldots, z_m)\end{aligned}$$

for a certain element f of G. Since we find that $(v, y_1, \ldots, y_n)$ and $(u, z_1, \ldots, z_m, v_m)$ commute, as we did in a similar case above, (3.9) and (1.3) give

$$c_4 = (v, y_1, \ldots, y_n; u, z_1, \ldots, z_m). \tag{3.10}$$

In the same way it can be shown that

$$c_1 = (u, y_1, \ldots, y_n; v, z_1, \ldots, z_m). \tag{3.11}$$

We rewrite (3.5) with the aid of (3.10) and (3.11):

$$(v, y_1, \ldots, y_n; u, z_1, \ldots, z_m) = (v, z_1, \ldots, z_m; u, y_1, \ldots, y_n). \tag{3.12}$$

Now suppose that $n > 1$ and that $y_n = v$ in (3.12). With the aid of the lemma of §2 we can express $(u, y_1, \ldots, y_{n-1}, v)$ as the product of commutators of the form $(v, y_1^*, \ldots, y_n^*)^{\pm 1}$. Then (3.2) shows that the right-hand side of (3.12) is 1, that is

$$(v, y_1, \ldots, y_{n-1}, v; u, z_1, \ldots, z_m) = 1. \tag{3.13}$$

Since $n \geqq m$ this states that in $G/Z(G_m)$ we have the law

$$(v, y_1, \ldots, y_{n-1}, v) = 1. \tag{3.14}$$

Apply Theorem 1. We see that $G/Z(G_m)$ is nilpotent of class $n+1$ or less, and so $G_m/Z(G_m)$ is certainly nilpotent of class no more than $n-m+1$. It follows that G_m is nilpotent of class $n-m+2$ or less.

This completes the proof of Theorem 2.

Note that our information about $G/Z(G_m)$ shows that G is an extension of an abelian group by a nilpotent group of class $n+1$ or less, so G is certainly soluble of length not exceeding $n+2$. We pass over the question whether the result of Theorem 2 is best possible.

4. The variety $(x, y; x, z) = 1$

The groups of this variety were excluded from the enunciation of Theorem 2. The metabelian groups, that is the variety with the law

$$(x, y; u, v) = 1$$

are closely related to our variety, perhaps more closely than might be expected; thus Theorem 3 shows that the groups of this variety are metabelian when they lack elements of order 2.

THEOREM 3. *Every group G of the variety associated with the law*

$$(x, y; x, z) = 1 \tag{4.1}$$

has the properties

(i) *G/Z is metabelian; and*

(ii) *G'' has exponent 2.*

PROOF. The law

$$\big((x, y)^u, (x, z)^v\big) = 1 \tag{4.2}$$

follows from (4.1) as (3.2) follows from (3.1). Several applications of (4.2) to (1.6) yield

$$(x, y, z)\,(y, z, x)\,(z, x, y) = 1\,. \tag{4.3}$$

Further (1.2) and (4.2) give

$$(x, y, z) = \big(z, (y, x)\big)^{(x, y)} = \big(z, (y, x)\big)\,.$$

We shall require two consequences of this equality:

$$(x, y; u, v) = (u, v; y, x)\,; \tag{4.4}$$

$$(x, y, z; u, v) = \big(z, (y, x); u, v\big)\,. \tag{4.5}$$

Next we may quote (3.12) with $m = n = 1$ as we did not use the assumption $n > 1$ in proving (3.12). Thus

$$(x, y; u, v) = (x, v; u, y)\,. \tag{4.6}$$

It is not difficult to deduce from (4.4) and (4.6) that the value of $(x, y; u, v)$ is unaltered under arbitrary permutation of x, y, u, v. Two consequences of this important fact are to be observed. Firstly

$$(x, y; u, v) = (u, v; x, y)$$

that is

$$(x, y; u, v)^2 = 1\,; \tag{4.7}$$

and secondly from (4.5) we have

$$(x, y, z; u, v) = (y, x, z; u, v)\,. \tag{4.8}$$

Next we need another relation, which comes from expansion of the law

$$(x, yz; u, v) = (x, v; u, yz) \tag{4.9}$$

derived from (4.6). Thus by (1.3) and (4.2)

$$\begin{aligned}(x, yz; u, v) &= \big((x, z)\,(x, y)^z, (u, v)\big)\\ &= (x, z; u, v)^{(x,y)^z}\,\big((x, y)^z, (u, v)\big)\\ &= (x, z; u, v)\,\big((x, y)\,(x, y, z), (u, v)\big)\\ &= (x, z; u, v)\,(x, y; u, v)^{(x,y,z)}\,(x, y, z; u, v)\\ &= (x, z; u, v)\,(x, y; u, v)\,(x, y, z; u, v);\end{aligned}$$

and similarly

$$(x, v; u, yz) = (x, v; u, y, z)\,(x, v; u, y)\,(x, v; u, z).$$

Therefore (4.9) becomes

$$\begin{aligned}&(x, z; u, v)\,(x, y; u, v)\,(x, y, z; u, v)\\ &\qquad = (x, v; u, y, z)\,(x, v; u, y)\,(x, v; u, z).\end{aligned}$$

We may rearrange factors in the last equality by (4.2), and cancel certain of them by (4.6). On using (4.6) again we have

$$(4.10)\qquad (u, v; x, y, z) = (u, y, v; x, z).$$

We are now in a position to make the final step in the argument. On applying (4.10), (4.8) and (4.10) again we obtain

$$(4.11)\qquad \begin{cases}(u, v; x, y, z) = (u, y, v; x, z)\\ \qquad\qquad\quad = (y, u, v; x, z)\\ \qquad\qquad\quad = (y, v; x, u, z).\end{cases}$$

Next we permute the elements $y, v, (x, u), z$ in the last commutator, use (4.8), then permute again:

$$(4.12)\qquad \begin{cases}(y, v; x, u, z) = (x, u, y; v, z)\\ \qquad\qquad\quad = (u, x, y; v, z)\\ \qquad\qquad\quad = (v, y; u, x, z).\end{cases}$$

But the argument, or result, of (4.11) gives

$$(4.13)\qquad (v, y; u, x, z) = (x, y; u, v, z).$$

Combining (4.11), (4.12) and (4.13) we obtain

$$(u, v; x, y, z) = (x, y; u, v, z)$$

which when variables are permuted establishes

$$(4.14)\qquad \big((x, y), z, (u, v)\big)\,\big(z, (u, v), (x, y)\big) = 1.$$

But from (4.3) we can derive the law

$$(4.15)\qquad \big((u, v), (x, y), z\big)\,\big((x, y), z, (u, v)\big)\,\big(z, (u, v), (x, y)\big) = 1.$$

On combining (4.14) with (4.15) we at once find

$$((u, v), (x, y), z) = 1; \tag{4.16}$$

that is, G'' lies in the centre of G. As G'' is abelian (4.7) shows that G'' has exponent 2.

This completes the proof of Theorem 3.

We apply the result to the 3-metabelian groups. These groups, discussed by B. H. NEUMANN in [3], are defined by the property that any subgroup generated by 3 elements is metabelian. It is therefore clear that every 3-metabelian group belongs to the variety (4.1).

COROLLARY. *Every 3-metabelian group G has G'' in the centre and of exponent* 2.

THEOREM 4. *Every group G of the variety associated with the law*

$$(x, y)^2 = 1 \tag{4.17}$$

has the properties

(i) *G'' lies in the centre; and*

(ii) *G' has exponent* 4.

PROOF. It will suffice to show that the variety (4.1) contains the variety (4.17), for (i) would follow at once from Theorem 3 and (ii) would follow from the fact that G' is generated by elements of order 2 and is nilpotent of class at most 2.

Now (1.8) gives

$$(y, x)\,(x, z) = (y z^{-1}, x)^z$$

identically, and in the variety (4.17) we deduce that

$$(y, x)\,(x, z)\,(y, x)\,(x, z) = 1$$

which is clearly equivalent to

$$(x, y; x, z) = 1$$

in this case.

This completes the proof of Theorem 4.

No one seems to have decided whether a group in which all commutators have finite, or bounded, order can have a non-periodic commutator subgroup.

That Theorems 3 and 4 are in a sense best possible results follows from the construction of a 3-metabelian non-metabelian group by NEUMANN in [3]. It is not difficult to verify that this example belongs to the variety (4.17) and so to the variety (4.1). More generally it can be proved that the 3-metabelian groups form the variety (4.1) — again, the details are omitted.

5. The variety $(x, y, z; x, w) = 1$

This variety is given by (3.1) with $m = 1$, $n = 2$, and so the conclusions of Theorem 2 are applicable; but more can be deduced when Theorem 3 also is used, as described below. Note that since the free nilpotent groups of

class 4 belong to this variety an exponent restriction on G'' cannot be expected.

THEOREM 5. *Every group G of the variety associated with the law*

$$(x, y, z; x, w) = 1 \tag{5.1}$$

has the properties

(i) *G/Z is metabelian; and*

(ii) *(G_2, G_1) has exponent 3.*

PROOF. The law

$$\big((x, y, z)^u, (x, w)^v\big) = 1 \tag{5.2}$$

follows from (5.1) as (3.2) follows from (3.1). We deduce from (1.1), (1.2) and (5.1) that

$$(x, y, z)^{-1} = \big(z, (x, y)\big) = (y, x, z)^{(x, y)} = (y, x, z);$$

in particular

$$(x, y; u, v) = (u, v; y, x) = (y, x; v, u) \tag{5.3}$$

and

$$(x, y, z; u, v) = (y, x, z; v, u). \tag{5.4}$$

Consider (1.5) with $a = w$, $b = (x, y)$, $c = (z, x)$:

$$\big(z, x, w^{-1}, (x, y)\big)^w \big(w, (y, x), (z, x)\big)^{(x, y)} \big(x, y, (x, z), w\big)^{(z, x)} = 1. \tag{5.5}$$

On applying (1.2) and (5.2) to the terms on the left-hand side we obtain

$$\big(x, y, (x, z), w\big) = 1. \tag{5.6}$$

As (5.6) states that G/Z belongs to the variety (4.1), the conclusions of Theorem 3 apply to G/Z. Therefore in G we have

$$\big((x, y; u, v)^2, z\big) = 1. \tag{5.7}$$

On using (1.3) and (5.2) we find that (5.7) becomes

$$\big((x, y), (u, v), z\big)^2 = 1. \tag{5.8}$$

Further, (4.3) shows that in G

$$(x, y, z)\,(y, z, x)\,(z, x, y) \in Z.$$

This fact with (5.2) gives

$$(y, z, x; x, w) = 1 \tag{5.9}$$

and

$$\big((x, y, z)\,(y, z, x)\,(z, x, y), (u, v)\big) = 1.$$

The last equality gives, when (1.3) and (5.2) are applied to it,

$$(x, y, z; u, v)\,(y, z, x; u, v)\,(z, x, y; u, v) = 1. \tag{5.10}$$

Next we quote (3.12) with $m=1$ and $n=2$:

$$(x, y, z; u, v) = (x, v; u, y, z). \tag{5.11}$$

We require a further relation, and it is obtained by expansion of the relation

$$(uv, (y, z); uv, w) = 1 \tag{5.12}$$

derived from (5.9). We may use the expansion of (3.3) with $m=m=1$; thus by (3.4) and (3.5)

$$(v, (y, z), (u, w)^v) = (v, w, (u, (y, z))^v). \tag{5.13}$$

After some reduction with (1.3), (5.2) and (5.9) we find

$$(v, (y, z), (u, w)) = ((v, w), (u, (y, z)));$$

and on using (5.3) this becomes

$$(y, z, v; u, w) = (y, z, u; w, v). \tag{5.14}$$

Now we shall use the earlier relations to develop (5.14). Applying (5.11) to both sides we obtain

$$(y, w; u, z, v) = (y, v; w, z, u). \tag{5.15}$$

Next apply (5.4) to both sides of (5.15):

$$(w, y; z, u, v) = (v, y; z, w, u). \tag{5.16}$$

Finally apply (5.11) to both sides of (5.16):

$$(w, u, v; z, y) = (v, w, u; z, y). \tag{5.17}$$

At this point (5.10) and (5.17) combine to give

$$(x, y, z; u, v)^3 = 1. \tag{5.18}$$

To complete the proof we apply (1.5) with $a=z$, $b=(u, v)$, $c=(y, x)$, and simplify with (5.2) and (1.2):

$$(y, x, z^{-1}; u, v)^z (z, (v, u), (y, x)) (u, v, (x, y), z) = 1,$$

$$(y, x, z^{-1}; u, v)^z (u, v, z; y, x) = (u, v, (x, y), z)^{-1}. \tag{5.19}$$

The factors on the left-hand side of (5.19) commute by (5.2) and their product has order 3 or 1 by (5.18). But the element on the right-hand side has order 2 or 1 by (5.8). Hence

$$(u, v, (x, y), z) = 1. \tag{5.20}$$

We conclude that $G'' \leqq Z(G)$; since G'' is abelian (5.18) shows that (G_2, G_1) has exponent 3. This completes the proof of Theorem 5.

COROLLARY 1. *Every group G whose 4-generator subgroups all belong to the variety*

$$(x, y, z; u, v) = 1 \tag{5.21}$$

has G'' central and (G_2, G_1) of exponent 3.

This is obvious. It can further be proved that the groups of the variety (5.1) are the only groups with the properties postulated in Corollary 1. A more interesting consequence of Theorem 5 refers to Theorems 3 and 4.

COROLLARY 2. *Every group G of the variety associated with the law* (4.1) *or the law* (4.17) *has* $(G_2, G_1) = 1$.

We saw that both these laws imply (4.2), which in turn implies (5.1), and so (G_2, G_1) has exponent 3. But we know from Theorems 3 and 4 that G'' has exponent 2. Therefore $(G_2, G_1) = 1$.

Corollary 2 can of course be proved more directly from the hypothesis of Theorem 3. We mention without proof that this result with the aid of (1.5) will give part (i) of Theorem 3.

References

[1] HALL, M.: The Theory of Groups. New York 1959.

[2] LEVI, F. W.: Groups in which the commutator relation satisfies certain algebraic conditions. J. Indian Math. Soc., N.s. **6**, 87–97 (1942).

[3] NEUMANN, B. H.: On a conjecture of Hanna Neumann. Proc. Glasgow Math. Assoc. **3**, 13–17 (1956).

Department of Pure Mathematics, The University, Sheffield 10

(Eingegangen am 3. Oktober 1960)

AVANN, S. P.
Math. Zeitschr. 76, 283–287 (1961)

Distributive properties in semi-modular lattices

By
S. P. AVANN

Projectivity and distributivity play an important role in the structure of lattices. In this study we are concerned with essentially distributive-like properties in sets of projective quotients in semi-modular lattices of finite length.

Two typical results are: a necessary and sufficient condition that a full equivalence class of projective quotients of equal length in a lower semi-modular lattice have a common lower transpose in the set is that numerators are unique; a merely sufficient condition is that the set of join-irreducibles contained in the numerator but not in the denominator of a quotient is invariant.

NOTATION. We shall use $\supset$, $\supseteq$, $>$ for proper inclusion, inclusion, and covering respectively between elements of the lattice $\mathscr{L}$ as well as between sets. $\delta[a]$ is rank or dimension of $a \in \mathscr{L}$. "Iff" means "if and only if".

$\mathscr{L}$ is a lower semi-modular or LSM lattice iff it satisfies

$$(\xi') \qquad b \cup c > b \quad \text{implies} \quad c > b \cap c.$$

$\mathscr{L}$ is upper semi-modular or USM iff it satisfies

$$(\xi) \qquad c > b \cap c \quad \text{implies} \quad b \cup c > b.$$

$\mathscr{L}$ is a lower semi-distributive or LSD (upper semi-distributive or USD) lattice if it is an LSM (USM) lattice in which all modular sublattices are distributive.

We designate a quotient lattice $\{v \in \mathscr{L} \mid a \supseteq v \supseteq b\}$ by a/b.

We write $a/b \supseteq_T c/d$ when a/b is merely an upper transpose of c/d: $a = b \cup c$ and $d = b \cap c$. When a/b and c/d are also of the same length: $\ln(a/b) = \delta[a] - \delta[b] = \delta[c] - \delta[d] = \ln(c/d)$, we write simply $a/b \supseteq c/d$ and call the transpose restricted. Restricted transposedness $\supseteq$ is a partial ordering relation on any full set of quotients of equal length; proper upper transposedness $\supset$, covering transposedness $>$, and the duals $\subseteq$, $\subset$, $<$ are defined naturally. When $\mathscr{L}$ is semi-modular, $Q_R = \{a_i/b_i\}$ will denote a full equivalence class with respect to restricted projectivity, which is the equivalence relation generated by restricted transposedness. For simplicity we shall hereafter call Q_R a restricted projective set. When $\mathscr{L}$ is modular, projective quotients are isomorphic, and the notion of restrictivity is redundant. We let

$$\mathscr{A} = \{a_i \mid a_i/b_i \in Q_R\} \quad \text{and} \quad \mathscr{B} = \{b_i \mid a_i/b_i \in Q_R\}.$$

Let P and P' be the partially ordered sets of join-irreducibles and of meet-irreducibles respectively of $\mathscr{L}$. We write $A=\{x\in P\mid x\leqq a\}$ and $A'=\{y\in P'\mid y\geqq a\}$; similarly $B, C, D, \ldots$ for $b, c, d, \ldots$. We shall use freely the following obvious properties: $A>B$ iff $a>b$; $a=\bigcup_{x\in A}x$; when $a/b\geqq_T c/d$, $A\geqq B+C$, $D=B\cdot C$, and $A-B\geqq C-B\cdot C=C-D$ where $+$, $\cdot$, and $-$ are point set sum, product, and difference. In the proof of Theorem 8 we shall use the dual properties of A' also.

Our results concern the following 11 conditions on a restricted projective set Q_R in a semi-modular lattice $\mathscr{L}$.

(1) Corresponding to each denominator $b_i\in\mathscr{B}$ is a unique numerator $a_i\in\mathscr{A}$.

(2) There exists $a_0\in\mathscr{A}$ such that $a_0\leqq a_i$ for all $a_i\in\mathscr{A}$.

(3) $\mathscr{A}$ is closed with respect to meet.

(4) There exists $a_0/b_0\in Q_R$ such that $a_0/b_0\leqq a_i/b_i$ for all $a_i/b_i\in Q_R$.

(5) $A_i-B_i=A_j-B_j$ for all a_i/b_i, a_j/b_j in Q_R.

(6) $\mathscr{A}\geqq\mathscr{B}$: As quotients $\mathscr{A}$ is an upper restricted transpose of $\mathscr{B}$

(1′)–(5′), the duals of (1)–(5).

We note that (6) is self dual.

For several proofs we need the following generalization of Dedekind's Transposition Principle.

LEMMA. *If $a/b>c/d$ and $a\geqq r>s\geqq c$ in a LSM lattice, then $b\geqq b\cap r>b\cap s\geqq d$ and $a/b\geqq r/b\cap r>s/b\cap s\geqq c/d$.*

PROOF. First, $a/b\geqq_T r/b\cap r$ and $c\cup(b\cap r)/b\cap r\geqq_T c/d$. Also $r\geqq c\cup(b\cap r)\geqq b\cap r$. Hence by a well-known extension of (ξ') [2; Lemma 1, p. 67] $\ln a/b\geqq\ln r/b\cap r\geqq\ln c\cup(b\cap r)/b\cap r\geqq\ln c/d=\ln a/b$. Therefore the lengths are equal and $r=c\cup(b\cap r)$. Thus $a/b\geqq r/b\cap r\geqq c/d$. Similarly, $r/b\cap r\geqq s/(b\cap r)\cap s=s/b\cap s\geqq c/d$. By ($\xi'$) $b\cap r>b\cap s$ so that $r/b\cap r>s/b\cap s$.

THEOREM 1. *In a LSM lattice* (1), (2), (3), *and* (4) *are equivalent for any restricted projective set Q_R.*

PROOF. (4) implies (1) and (2), and (3) implies (2) trivially. Next, (4) implies (3). For consider $a_r/b_r\geqq a_0/b_0\leqq a_t/b_t$. Since $a_r\geqq a_r\cap a_t\geqq a_0$, the lemma yields $a_r/b_r\geqq a_r\cap a_t/(a_r\cap a_t)\cap b_r=a_r\cap a_t/b_r\cap b_t\geqq a_0/b_0$. Whence $a_r\cap a_t/b_r\cap b_t\in Q_R$ and $a_r\cap a_t\in\mathscr{A}$ as desired. Given (2), we shall prove (4). Let b_0 be a denominator of a_0. By transitivity $a_0/b_0\leqq a_r/b_r\leqq a_t/b_t$ implies $a_0/b_0\leqq a_t/b_t$. Now consider the case $a_0/b_0\leqq a_r/b_r\geqq a_t/b_t$. We have $a_r=a_t\cup b_r\geqq a_t\geqq a_0$, the latter by hypothesis. By the lemma $a_0/b_0\leqq a_t/b_r\cap a_t=a_t/b_t\leqq a_r/b_r$. For arbitrary $a/b\in Q_R$ there exists a finite sequence $a_0/b_0\leqq a_n/b_n\sim a_t/b_t\sim\cdots\sim a/b$ where $\sim$ is restricted transposition $\geqq$ or $\leqq$. Applying finite induction, the above two properties yield $a_0/b_0\leqq a/b$ for all $a/b\in Q_R$. The proof of (1) implies (4) follows the pattern of the preceding. Let b_0 be any minimal denominator and a_0 its unique numerator. Again $a_0/b_0\leqq a_r/b_r\leqq a_t/b_t$ implies $a_0/b_0\leqq a_t/b_t$. Hence we need only show that $a_0/b_0\leqq a_r/b_r\geqq a_t/b_t$ implies $a_0/b_0\leqq a_t/b_t$ to obtain (4). Accordingly, we assume $a_t\geqq a_0$. First, $a_t\not\leqq a_0$; otherwise the lemma gives

$a_t/b_t \subset a_r/b_0 \subseteq a_r/b_r$ contradicting the minimality of b_0. Hence there exists a connected chain $a_0 < a_1 < \cdots < a_n \equiv a_0 \cup a_t$ with $n \geqq 1$ and an element c such that $a_t \subseteq c < a_n$ with $c \not\supseteq a_0$. Let $b_i = b_r \cap a_i$ and $e_i = c \cap a_i$ $(i = 0, 1, 2, \ldots, n)$. By the lemma $b_0 < b_1 < \cdots < b_n \subseteq b_r$ and $a_0/b_0 < a_1/b_1 < \cdots < a_n/b_n \subseteq a_r/b_r$. We show by an induction on i that $b_i \subseteq e_i < a_i$ $(i = 0, 1, 2, \ldots, n)$ and that $e_0 < e_1 < \cdots < e_{n-1} < e_n = c$. For $i = 0$ we have $e_0 < a_0$ by (ξ'). Moreover $b_0 \subseteq e_0$; otherwise $b_0 > b_0 \cap e$ by (ξ') and $a_0/b_0 > e_0/b_0 \cap e_0$, contradicting the minimality of b_0. Assume true for $0 \leqq i < n$ and consider the case of $i+1$. Now by the induction hypothesis $a_{i+1} > a_i > e_i \geqq b_i$. Since $c \not\supseteq a_0$, $c \not\supseteq a_{i+1}$, $a_n = c \cup a_{i+1} > c$, and by (ξ') $a_{i+1} > e_{i+1} \geqq e_i$. Indeed, $e_{i+1} > e_i$ by the Jordan Dedekind chain condition. Suppose $e_{i+1} \not\supseteq b_{i+1}$. Then $b_i \subseteq e_{i+1} \cap b_{i+1} \subset b_{i+1}$ forcing $b_{i+1} > b_i = e_{i+1} \cap b_{i+1}$ and $a_{i+1} = e_{i+1} \cap b_{i+1} > e_{i+1}$. Thus $a_{i+1}/b_{i+1} > e_{i+1}/b_i \in Q_R$. But the unique numerator of b_i is a_i, which is distinct from e_{i+1}, since $c \not\supseteq a_i$. Thus $e_{i+1} \geqq b_{i+1}$ and our induction on i is complete. Hence $a_n > c = e_n \geqq b_n$ and $b_n = b_n \cap c$. Applying the lemma to $a_r \geqq a_n > c \geqq a_t$ we obtain $b_r \geqq b_n > b_r \cap c = b_r \cap a_n \cap c = b_n \cap c$, a contradiction. Hence our initial assumption $a_t \not\supseteq a_0$ is false. Thus $a_r \geqq a_t \geqq a_0$ and by the lemma $a_r/b_r \geqq a_t/b_t \geqq a_0/b_0$ as desired. Therefore (1) implies (4) and the proof of the entire theorem is complete.

REMARK. In a *USM* or *USD* lattice *USD* counterexamples show that (1) does not imply (2), (3), or (4) and (3) is not implied by (1), (2), or (4) in Q_R. It .s easy to show that (2) ,mplies (4). All the other implications are trivial.

THEOREM 2. *In any restricted projective set Q_R of an LSM lattice* (5) *implies* (1), (2), (3), *and* (4).

PROOF. If (5) holds, a given denominator $b \in \mathscr{B}$ uniquely determines B, A, and $a = \bigcup_{x \in A} x$, which is condition (1). Hence also (2), (3), and (4) hold by Theorem (1).

The converse of Theorem 2 is not true.

THEOREM 3. *In a LSD lattice $\mathscr{L}$* (1)—(5) *all hold for any restricted projective set Q_R.*

PROOF. By Theorem 2 it suffices to prove (5) holds in Q_R. Let $a/b \geqq c/d$ be an arbitrary transposed pair in Q_R; then $\delta[a] - \delta[b] = \delta(c] - \delta[d]$. But it is known [*3* or *1*, Theorem 5.5 (H) iff (B)] that $\delta[a]$ is the order of A for all $a \in \mathscr{L}$. Since always $A \geqq B + C$ and $A - B \geqq C - D$, it then follows that $A - B = C - D$.

COROLLARY. *A LSM lattice is LSD iff for each set Q_R any one of* (1)—(5) *hold.*

PROOF. Necessity by Theorem 3. If $\mathscr{L}$ is not *LSD*, then there exists a non-distributive modular sublattice, another of order 5, and a third of order 5 [*1*; Corollary 4.4.3] which is dense in $\mathscr{L}$: $c > a_1, a_2, a_3 > b$. Then a_1/b, a_2/b, and a_3/b belong to the same restricted set Q_R of prime quotients. Hence (1) is false, and each of (2)—(5) are likewise false for Q_R by the contrapositives of Theorems 1 and 2.

THEOREM 4. *For any projective set Q of quotients in a modular lattice $\mathscr{L}$ each pair of conditions, one from* (1), (2), (3), (4) *and one from their duals* (1'), (2'), (3'), (4') *is equivalent to* (6).

PROOF. By the dual of Theorem 1 (1'), (2'), (3'), (4') are equivalent in an *USM* lattice, hence in $\mathscr{L}$ also. Clearly (4) and (4'): $a_0/b_0 \leqq a_i/b_i \leqq a_1/b_1$ for all $a_i/b_i \in Q$, is equivalent to $\mathscr{A} = a_1/a_0 \geqq b_1/b_0 = \mathscr{B}$.

THEOREM 5. *For any projective set Q of quotients that are complemented sublattices in a modular lattice $\mathscr{L}$, each of the* 11 *conditions* (1)—(5), (1')—(5'), (6) *are equivalent.*

PROOF. In view of Theorems 1, 2, their duals, and 4 it will suffice to prove that (1) implies (1') and (5). For then by duality (1') implies (1) and (5'). Given (1), let b_2 and b_3 be two denominators of an arbitrary numerator a_2. Let c_2' be a complement of $c_2 \equiv b_2 \cup b_3$ in a_2/b_2 and c_3' be a complement of c_2 in a_2/b_3. Then $c_2' \cup b_3 = c_2' \cup b_2 \cup b_3 = c_2' \cup c_2 = a_2$ and $c_2' \cap b_3 = c_2' \cap c_2 \cap b_3 = b_2 \cap b_3$. We therefore have $a_2/b_3 \geqq c_2'/b_2 \cap b_3 \in Q$. By symmetry $a_2/b_2 \geqq c_3'/b_2 \cap b_3 \in Q$. By uniqueness of numerators $c_2' = c_3'$. It follows that $b_2 = c_2' \cap c_2 = c_3' \cap c_2 = b_3$. Hence (1') holds and by Theorem 1 and its dual (4) and (4') hold in particular: $a_1/b_1 \geqq a_i/b_i \geqq a_0/b_0$ for a_1/b_1, a_0/b_0, and all $a_i/b_i \in Q$. Assume (5) is false; then since $A_1 - B_1 \geqq A_i - B_i \geqq A_0 - B_0$, it must follow that $A_1 - B_1 > A_0 - B_0$. Particularly, by considering a connected descending chain from a_1 to a_0, there exists a/b, $c/d \in Q$ such that $a_1/b_1 \geqq a/b > c/d \geqq a_0/b_0$ and $A - B > C - D$. Consider $x \in (A - B) - (C - D)$: $x \leqq a$, $x \not\leqq b$, $x \not\leqq c$. The join-irreducible x covers exactly one element x_0. Now $a \geqq x \cup c > c < a$ requires $a = x \cup c > c$. Hence by (ξ') $x > x \cap c = x_0$ necessarily. By the modular law $c \cap (x \cup d) = (c \cap x) \cup d = x_0 \cup d$. Also $c \cup (x \cup d) = c \cup x = a$. Application of ($\xi'$) to $b \cap (x_0 \cup d) = b \cap c \cap (x_0 \cup d) = d \cap (x_0 \cup d) = d < b$ yields $b \cup (x_0 \cup d) = x_0 \cup b > x_0 \cup d$. Again by ($\xi'$) $c \cup (x_0 \cup b) = x_0 \cup a = a > c$ leads to $x_0 \cup b > c \cap (x_0 \cup b) \geqq x_0 \cup d$, requiring $c \cap (x_0 \cup b) = x_0 \cup d$. The prime quotient b/d is trivially complemented, while c/d is complemented by hypothesis. Hence $a/d = b \cup c/d$ is complemented [2; Theorem 6, p. 105]. Its quotient sublattice $a/x_0 \cup d$ is therefore likewise complemented. We have shown that in the latter c has two relative complements $x \cup d$ and $x_0 \cup b$, which we shall show are distinct. Since $b \cap x < x$, then $b \cap x \leqq x_0$ and the modular law gives $x_0 = x_0 \cup (b \cap x) = (x_0 \cup b) \cap x < x$. Therefore $x_0 \cup b \not\geqq x \leqq x \cup d \neq x_0 \cup b$. Hence by a theorem of VON NEUMANN [2; Theorem 10, p. 124], c is not in the center of $a/x_0 \cup d$. Neither then is the relative complement $x_0 \cup b$ in the center, and by the same theorem $x_0 \cup b$ has another complement e in $a/x_0 \cup d$ distinct from c. But $e \cup b = e \cup x_0 \cup b = e \cup a = a$ and $e \cap b = e \cap (x_0 \cup b) \cap b = (x_0 \cup d) \cap b = (x_0 \cup d) \cap c \cap b = (x_0 \cup d) \cap d = d$. Hence $a/b > e/d \in Q$ making c and e distinct numerators of d. This contradiction of our hypothesis (1) means our assumption is wrong. Thus (1) implies (5).

THEOREM 6. *In each projective set Q of a distributive lattice all* 11 *conditions* (1)—(5), (1')—(5'), (6) *hold.*

PROOF. Apply Theorem 3, its dual, and Theorem 4.

COROLLARY. *A modular lattice is distributive iff for each set Q any one of the* 11 *conditions holds.*

PROOF. Necessity by Theorem 6. A non-distributive modular lattice $\mathscr{L}$ contains a non-distributive modular sublattice of order 5 that is dense in $\mathscr{L}$: $c > a_1, a_2, a_3 > b$. Here a_1, a_2, a_3 are distinct numerators of b in the set Q of (prime) complemented quotients to which a_1/b, a_2/b, a_3/b belong. Hence in Q (1) is false, and all the other 10 conditions are likewise false by the contrapositive of Theorem 5.

DEFINITION. We call the elements b, c a distributive pair in a lattice $\mathscr{L}$ iff all six distributive laws hold among b, c, and g for all $g \in \mathscr{L}$. When $\mathscr{L}$ is modular, one distributive law implies all six.

THEOREM 7. *In a modular lattice b, c are a distributive pair if and only if $A = B + C$ where $a = b \cup c$.*

PROOF. Let b, c be a distributive pair. $A \geqq B + C$. For arbitrary join-irreducible $x \in A$ $x = x \cap a = x \cap (b \cup c) = (x \cap b) \cup (x \cap c)$. Hence $x = x \cap b$ or $x = x \cap c$. I.e. $x \leqq b$ or $x \leqq c$, so that $x \in (B + C)$ and $A = B + C$. Conversely, suppose $A = B + C$. For arbitrary $e \in \mathscr{L}$ $E \cdot A = E \cdot (B + C) = E \cdot B + E \cdot C \leqq G$, where $g = (e \cap b) \cup (e \cap c)$. Hence $e \cap a \leqq g$. But $e \cap a \geqq g$ always. Hence equality subsists.

THEOREM 8. *In a projective set Q of a modular lattice* (5) *and* (5′) *are equivalent (and either implies all the other ten conditions).*

PROOF. Given (5). Then (4) holds by Theorem 2. Given an arbitrary numerator $a_i \in \mathscr{A}$; a_i determines A_i, which in turn determines B_i by hypothesis and $b_i = \bigcup_{x \in B_i} x$, all uniquely. Thus (1′), and by the dual of Theorem 1, (4′) hold true. Hence $a_0/b_0 \leqq a_i/b_i \leqq a_1/b_1$ for a_0/b_0, a_1/b_1 and all $a_i/b_i \in Q$. We then have $A_1 - B_1 = A_0 - B_0 = A_0 - B_1 \cdot A_0$ and $A_1 = A_0 + B_1$. By Theorem 7, a_0 and b_1 are a distributive pair. By the dual of the converse $B_0' = B_1' + A_0'$ and $B_0' - A_0' = B_1' - B_1' \cdot A_0' = B_1' - A_1'$. But $B_0' - A_0' \geqq B_i' - A_i' \geqq B_1' - A_1'$ for all $a_i/b_i \in Q$ then entails equality, which is condition (5′). Dually, (5′) implies(5).

References

[1] AVANN, S. P.: Application of the join-irreducible excess function to semi-modular lattices. Math. Annalen. **142**, 345—354 (1961).

[2] BIRKHOFF, G.: Lattice theory, rev. ed. Amer. Math. Soc. Colloquium Publications, vol. 25, 1948.

[3] DILWORTH, R. P.: Lattices with unique irreducible decompositions. Ann. Math. **41**, 771—777 (1940).

Department of Mathematics, University of Washington, Seattle 5

(Eingegangen am 25. Oktober 1960)

ISHIGURO, K.
Math. Zeitschr. 76, 288–294 (1961)

Zur Gibbsschen Erscheinung für das Kreisverfahren

Von

KAZUO ISHIGURO

1. Bei den folgenden Untersuchungen studieren wir die Gibbssche Erscheinung der Fourier-Reihe $\sum_{\nu=1}^{\infty} \frac{\sin \nu t}{\nu}$. Das Folgende ist schon lange bekannt [1]: Wenn $nt_n \to \tau$ und $t_n \to +0$ für $n \to \infty$, dann strebt dafür

$$s_n(t_n) \to \int_0^{\tau} \frac{\sin v}{v}\, dv, \qquad 0 \leqq \tau \leqq +\infty,$$

wobei

$$s_n(t) = \sum_{\nu=1}^{n} \frac{\sin \nu t}{\nu}.$$

O. Szász [2] bewies den folgenden

SATZ 1. *Für das Eulersche Mittel $\sigma_n(t)$ der Fourier-Reihe $\sum_{\nu=1}^{\infty} \frac{\sin \nu t}{\nu}$ gilt: Wenn $nt_n \to \tau$, $0 \leqq \tau \leqq +\infty$, und $nt_n^2 \to 0$ für $t_n \to +0$, dann strebt dafür*

$$\sigma_n(t_n) \to \int_0^{\tau r} \frac{\sin v}{v}\, dv,$$

wobei

$$\sigma_{n,r}(t) = \sigma_n(t) = \sum_{\nu=0}^{n} \binom{n}{\nu} r^{\nu} (1-r)^{n-\nu} s_{\nu}(t), \qquad \text{(vgl. [3], [4], [9])}.$$

Die $\sigma_n(t)$ sind abhängig vom Parameter r, von dem wir $0 < r \leqq 1$ annehmen. Bei Annahme $r = 1$ reduziert sich die obige Definition auf die der gewöhnlichen Konvergenz. Wir bezeichnen dies Limitierungsverfahren als das (ε, r)-Verfahren.

Wir beweisen hier den folgenden

SATZ 2. *Für das Kreismittel $\sigma_n^*(t)$ der Fourier-Reihe $\sum_{\nu=1}^{\infty} \frac{\sin \nu t}{\nu}$ gilt: Unter denselben Bedingungen über t_n wie in Satz 1 strebt*

$$\sigma_n^*(t_n) \to \int_0^{\tau/r} \frac{\sin v}{v}\, dv \quad \text{für} \quad n \to \infty,$$

wobei

$$\sigma^*_{n,r}(t) = \sigma^*_n(t) = \sum_{\nu=n}^{\infty} \binom{\nu}{n} r^{n+1}(1-r)^{\nu-n} s_\nu(t), \quad \text{(vgl. [4], [9])}.$$

Die $\sigma^*_n(t)$ sind abhängig vom Parameter r, von dem wir $0<r\leqq 1$ annehmen. Bei Annahme $r=1$ reduziert sich die obige Definition auf die der gewöhnlichen Konvergenz. Wir bezeichnen dies Limitierungsverfahren als das (γ, r)-Verfahren.

2. Zum Beweis von Satz 2 benützen wir die Methode von O. Szász [2].

Da $s_0=0$, $s_n=s_n(t)=\sum_{\nu=1}^{n} \frac{\sin \nu t}{\nu} = -\frac{t}{2} + \int_0^t \frac{\sin(n+\frac{1}{2})x}{2\sin\frac{1}{2}x}\,dx$, so ist

$$\sigma^*_n(t) = \sum_{\nu=n}^{\infty} \binom{\nu}{n} r^{n+1}(1-r)^{\nu-n} s_\nu(t)$$

$$= -\frac{t}{2} + \int_0^t \frac{1}{2\sin\frac{1}{2}x} \sum_{\nu=n}^{\infty} \binom{\nu}{n} r^{n+1}(1-r)^{\nu-n} \sin\left(\nu+\frac{1}{2}\right)x\,dx,$$

wo die Umkehrung von zwei Operationen wegen der gleichmäßigen Konvergenz der Summierung erlaubt ist. Andererseits folgt aus

$$\sum_{\nu=n}^{\infty} \binom{\nu}{n} (1-r)^{\nu-n} e^{i(\nu-n)x} = \frac{1}{\{1-(1-r)e^{ix}\}^{n+1}},$$

daß

$$\sigma^*_n(t) = -\frac{t}{2} + \frac{1}{2}\Im \int_0^t \frac{r^{n+1} e^{i(n+\frac{1}{2})x}}{\sin\frac{1}{2}x\,\{1-e^{ix}+re^{ix}\}^{n+1}}\,dx.$$

Setzen wir nun

$$\frac{1}{1-e^{ix}+re^{ix}} = \varrho e^{i\alpha},$$

oder

$$\text{(1)} \qquad \begin{cases} \frac{1}{\varrho}\cos\alpha = 1-\cos x + r\cos x \\ \frac{1}{\varrho}\sin\alpha = \sin x - r\sin x, \end{cases}$$

so ist

$$\text{(2)} \qquad \begin{cases} \frac{1}{\varrho^2} = r^2 + 2(1-r)(1-\cos x) \\ \frac{1}{r^2\varrho^2} \geqq 1 \end{cases}$$

und weiter

$$\text{(3)} \qquad 0 \leqq r\varrho \leqq 1.$$

Demnach ist

$$(4)\quad \begin{cases} \sigma_n^*(t)+\frac{t}{2}=\frac{1}{2}\int\limits_0^t \frac{1}{\sin\frac{1}{2}x}\, r^{n+1}\varrho^{n+1}\sin\left\{(n+1)\alpha+\left(n+\frac{1}{2}\right)x\right\}dx \\ \qquad =\frac{1}{2}\int\limits_0^t \cot\frac{1}{2}x\,(r\varrho)^{n+1}\sin\{(n+1)\alpha+nx\}\,dx+ \\ \qquad\quad +\frac{1}{2}\int\limits_0^t (r\varrho)^{n+1}\cos\{(n+1)\alpha+nx\}\,dx=\frac{1}{2}(I+J). \end{cases}$$

Nun ist wegen (3)

$$|J|=\left|\int\limits_0^t (r\varrho)^{n+1}\cos\{(n+1)\alpha+nx\}\,dx\right|\leqq t,$$

so daß $J=\eta t$, $(|\eta|<1)$, und

$$\sigma_n^*(t)+\frac{1-\eta}{2}t=\frac{1}{2}\int\limits_0^t (r\varrho)^{n+1}\cot\frac{1}{2}x\sin\{(n+1)\alpha+nx\}\,dx.$$

Andererseits ist

$$(5)\qquad 1-(r\varrho)^{n+1}=(1-r\varrho)\sum_{\nu=0}^{n}(r\varrho)^{\nu}\leqq(n+1)(1-r\varrho),$$

und aus (2) folgt

$$1-(r\varrho)^2=\frac{2(1-r)(1-\cos x)}{r^2+2(1-r)(1-\cos x)}.$$

Da $1-\cos x\leqq\frac{x^2}{2}$ für kleines x, gilt

$$1-(r\varrho)^2\leqq\frac{1-r}{r^2}x^2 \quad\text{und}\quad 1-r\varrho\leqq\frac{1-r}{r^2}x^2.$$

Nach (5) ist daher

$$1-(r\varrho)^{n+1}\leqq\frac{1-r}{r^2}(n+1)x^2,$$

oder auch

$$1-(r\varrho)^{n+1}=\lambda(n+1)x^2,\quad\text{wobei}\quad 0\leqq\lambda\leqq\frac{1-r}{r^2}.$$

Hieraus folgt

$$I=\int\limits_0^t (r\varrho)^{n+1}\cot\tfrac{1}{2}x\sin\{(n+1)\alpha+nx\}\,dx$$

$$=\int\limits_0^t \cot\tfrac{1}{2}x\sin\{(n+1)\alpha+nx\}\,dx-(n+1)\int\limits_0^t \lambda x^2\cot\tfrac{1}{2}x\sin\{(n+1)\alpha+nx\}\,dx$$

$$=I_1-I_2.$$

Hier ist

$$|I_2| = (n+1)\left|\int_0^t \lambda x^2 \cot\frac{1}{2}x \sin\{(n+1)\alpha + nx\}\,dx\right|$$

$$\leqq (n+1)\frac{1-r}{r^2}\int_0^t \left|x^2\cot\frac{1}{2}x\right| dx = O(nt^2) = o(1)$$

für $n\to\infty$, wegen der Voraussetzung. Um I_1 abzuschätzen, benützen wir den folgenden

HILFSSATZ (vgl. [5]). *Für kleines x ist*

$$\alpha = sx + O(x^3), \quad \textit{wobei} \quad s = \frac{1-r}{r}.$$

BEWEIS. Aus (1) folgt

$$\tan\alpha = \frac{(1-r)\sin x}{1-(1-r)\cos x} = \frac{(1-r)\sin x}{r+(1-r)(1-\cos x)}$$

$$= \frac{s\sin x}{1+2s\sin^2\frac{x}{2}}.$$

Andererseits ist für kleines x

$$\text{si}\quad r = x + Ax^3, \quad |A| \leqq 2,$$

und

$$\frac{1}{1+x} = 1 + Bx, \quad |B| \leqq 1, \quad \text{für} \quad x \geqq 0$$

Daher ist

$$\frac{s\sin x}{1+2s\sin^2\frac{x}{2}} = sx + O(x^3).$$

Weiter, da

$$\tan^{-1}x = x + Cx^3, \quad |C| \leqq 4 \quad \text{für kleines } x,$$

so folgt

$$\alpha = \tan^{-1}\frac{(1-r)\sin x}{1-(1-r)\cos x} = sx + O(x^3),$$

womit der Hilfssatz bewiesen ist.

Wegen $\alpha = sx + \mu x^3$, $|\mu| \leqq K$, erhalten wir

$$I_1 = \int_0^t \cot\frac{1}{2}x \sin\{(n+1)\alpha + nx\}\,dx$$

$$= \int_0^t \cot\frac{1}{2}x \sin\left\{sx + \frac{nx}{r} + (n+1)\mu x^3\right\} dx$$

$$= \int_0^t \cot\frac{1}{2}x \sin sx \cos\left\{\frac{nx}{r} + (n+1)\mu x^3\right\} dx +$$

$$+ \int_0^t \cot\frac{1}{2}x \cos sx \sin\left\{\frac{nx}{r} + (n+1)\mu x^3\right\} dx = I_{11} + I_{12}.$$

Hier ist $|I_{11}| = O(t) = o(1)$ für $t \to 0$. Zweitens gilt

$$I_{12} = \int_0^t \cot\frac{1}{2}x \cos s\,x \sin\left\{\frac{n\,x}{r} + (n+1)\,\mu\,x^3\right\} dx$$

$$= \int_0^t \cot\frac{1}{2}x \cos s\,x \sin\frac{n\,x}{r} \cos(n+1)\,\mu\,x^3\,dx$$

$$+ \int_0^t \cot\frac{1}{2}x \cos s\,x \cos\frac{n\,x}{r} \sin(n+1)\,\mu\,x^3\,dx = I_{121} + I_{122},$$

wobei

$$|I_{122}| \leqq (n+1)\,K \int_0^t \left|\cot\frac{1}{2}x\right| x^3\,dx = O(n\,t^3) = o(1),$$

$$I_{121} = 2\int_0^t \frac{\sin\frac{n}{r}x}{x}\,dx - \int_0^t \sin\frac{n\,x}{r}\left\{\frac{2}{x} - \frac{\cos\frac{1}{2}x\cos s\,x\cos(n+1)\,\mu\,x^3}{\sin\frac{1}{2}x}\right\} dx.$$

Hierbei ist

$$2\sin\tfrac{1}{2}x - x\cos\tfrac{1}{2}x\cos s\,x\cos(n+1)\,\mu\,x^3$$

$$= 2\sin\tfrac{1}{2}x - x\cos\tfrac{1}{2}x + x\cos\tfrac{1}{2}x\{1 - \cos s\,x\cos(n+1)\,\mu\,x^3\},$$

und weiter erhalten wir durch den Mittelwertsatz

$$0 \leqq 2\sin\tfrac{1}{2}x - x\cos\tfrac{1}{2}x < \tfrac{1}{4}x^3,$$

und

$$\int_0^t \sin\frac{n}{r}x\,\frac{2\sin\frac{1}{2}x - x\cos\frac{1}{2}x}{x\sin\frac{1}{2}x}\,dx = O\left(\int_0^t x\,dx\right) = o(1).$$

Weiter, da

$$1 - \cos s\,x\cos(n+1)\,\mu\,x^3$$

$$= 1 - \cos(n+1)\,\mu\,x^3 - \cos(n+1)\,\mu\,x^3(\cos s\,x - 1)$$

$$= O(n^2x^6) + O(x^2),$$

gilt

$$\int_0^t \sin\frac{n}{r}x\cot\frac{1}{2}x\{1 - \cos s\,x\cos(n+1)\,\mu\,x^3\}\,dx$$

$$= O\left(\int_0^t n^2x^5\,dx\right) + O\left(\int_0^t x\,dx\right)$$

$$= O(n^2t^6) + O(t^2) = o(1).$$

Daher erhalten wir

$$2\sigma_n^*(t_n) - (1-\eta)\, t_n = 2\int_0^{t_n} \frac{\sin\frac{n}{r}x}{x}\,dx + o(1)$$

$$= 2\int_0^{n t_n/r} \frac{\sin v}{v}\,dv + o(1),$$

so daß

$$\sigma_n^*(t_n) \to \int_0^{\pi/r} \frac{\sin v}{v}\,dv \quad \text{für} \quad n \to \infty.$$

Damit ist der Satz bewiesen.

3. W. Meyer-König [6] und P. Vermes [10] haben die Verwandtschaft zwischen dem (ε, r)-Verfahren und dem (γ, r)-Verfahren untersucht und Sätze über den Vergleich dieser Verfahren aufgestellt. Andererseits haben L. Lorch [7] und A. Livingston [8] die Lebesgueschen Konstanten für das (ε, r)-Verfahren der Fourier-Reihe angegeben, d.h.

$$L(n; r) = \frac{2}{\pi^2}\log\left|\frac{2nr}{1-r}\right| + A + o(1),$$

mit

$$A = -\frac{2C}{\pi^2} + \frac{2}{\pi}\int_0^1 \frac{\sin u}{u}\,du - \frac{2}{\pi}\int_0^\infty \left\{\frac{2}{\pi} - |\sin u|\right\}\frac{du}{u}, \tag{6}$$

wobei C die Euler-Mascheronische Konstante ist. Der Autor hat auch die Lebesgueschen Konstanten für das (γ, r)-Verfahren angegeben [5], d.h.

$$L^*(n; r) = \frac{2}{\pi^2}\log\left|\frac{2n}{1-r}\right| + A + o(1),$$

wobei A durch (6) gegeben ist.

Wenn wir das r in $L(n; r)$ durch $1/r$ ersetzen, so folgt $L^*(n; r)$, d.h. $L\left(n; \frac{1}{r}\right) = L^*(n; r)$. Wie wir schon gesehen haben, wenn wir in der Formel $\int_0^{\pi r} \frac{\sin v}{v}\,dv$ der Gibbsschen Erscheinung des (ε, r)-Verfahrens $\frac{1}{r}$ an Stelle des r setzen, so erhalten wir die Formel $\int_0^{\pi/r} \frac{\sin v}{v}\,dv$ von derjenigen des (γ, r)-Verfahrens. Diese Beziehung $r \leftrightarrow \frac{1}{r}$ zwischen (ε, r)-Verfahren und (γ, r)-Verfahren erscheint uns interessant.

Zum Schluß möchte der Autor seinen herzlichen Dank an die Herren K. Kunugi, S. Izumi und G. Pickert für ihre Beratung und Anregung ausdrücken.

Literatur

[1] ZYGMUND, A.: Trigonometric series. I. Cambridge 1959.
[2] SZÁSZ, O.: On the Gibbs' phenomenon for Euler means. Acta Univ. Szeged **12**, part B, 107–111 (1950).
[3] KNOPP, K.: Theorie und Anwendung der unendlichen Reihen. Berlin 1931.
[4] HARDY, G. H.: Divergent series. Oxford 1949.
[5] ISHIGURO, K.: The Lebesgue constants for (γ, r) summation of Fourier series. Proc. Japan Acad. **36**, 470–474 (1960).
[6] MEYER-KÖNIG, W.: Untersuchungen über einige verwandte Limitierungsverfahren. Math. Z. **52**, 257–307 (1950).
[7] LORCH, L.: The Lebesgue constants for $(E, 1)$ summation of Fourier series. Duke Math. J. **19**, 45–50 (1952).
[8] LIVINGSTON, A. E.: The Lebesgue constants for Euler (E, p) summation of Fourier series. Duke Math. J. **21**, 309–313 (1954).
[9] ZELLER, K.: Theorie der Limitierungsverfahren. Berlin 1958.
[10] VERMES, P.: Series to series transformations and analytic continuation by matrix methods. Amer. J. Math. **71**, 541–562 (1949).

Hokkaido-Universität, Sapporo (Japan)

(Eingegangen am 7. November 1960)

AHMAD, Z. U.
Math. Zeitschr. 76, 295–310 (1961)

On the absolute Cesàro summability factors of infinite series

By
Z. U. AHMAD

1.1. Definition. Let $\sum a_n$ be a given infinite series, and let s_n^ϱ and t_n^ϱ denote the n-th Cesàro means of order ϱ ($\varrho > -1$) of the sequences $\{s_n\}$ and $\{n\,a_n\}$ respectively, where s_n is the n-th partial sum of $\sum a_n$. The series $\sum a_n$ is said to be absolutely summable (C, ϱ), or summable $|C, \varrho|$, if the sequence $\{s_n^\varrho\}$ is of bounded variation, that is, the infinite series $\sum |s_n^\varrho - s_{n-1}^\varrho|$ is convergent [1]).

1.2. We shall require the following definitions and identities.

$$t_n^\varrho = n\,(s_n^\varrho - s_{n-1}^\varrho)\ ^{2}), \tag{1.2.1}$$

$$t_n^\varrho = \frac{1}{A_n^\varrho} \sum_{\nu=1}^{n} A_{n-\nu}^{\varrho-1}\,\nu\,a_\nu, \tag{1.2.2}$$

where

$$\sum_{n=0}^{\infty} A_n^\varrho x^n = (1-x)^{-\varrho-1} \qquad (|x|<1); \tag{1.2.3}$$

and, by definition [3]),

$$A_{-1}^\varrho = 0; \quad A_0^{-1} = 1; \quad A_n^{-\varrho} = 0 \,(n \geqq \varrho;\ \varrho = 1, 2, \ldots), \tag{1.2.4}$$

$$A_n^\varrho = \begin{cases} \binom{n+\varrho}{n} & (\varrho > -1) \\ (-1)^n \binom{-\varrho-1}{n} & (\varrho \leqq -1); \end{cases} \tag{1.2.5}$$

$$A_n^\varrho = \Gamma(n+\varrho+1)\,\{\Gamma(n+1)\,\Gamma(\varrho+1)\}^{-1} \qquad (\varrho \neq -1, -2, \ldots) \tag{1.2.6}$$
$$\sim n^\varrho \{\Gamma(\varrho+1)\}^{-1}.$$

For any sequence $\{\varepsilon_n\}$, we write

$$\Delta^0 \varepsilon_n = \varepsilon_n, \quad \Delta\,\varepsilon_n = \Delta^1 \varepsilon_n = \varepsilon_n - \varepsilon_{n+1}, \tag{1.2.7}$$

and

$$\Delta^\varrho \varepsilon_n = \sum_{\nu=0}^{\infty} A_\nu^{-\varrho-1}\,\varepsilon_{\nu+n}, \tag{1.2.8}$$

provided the series on the right converges.

[1]) FEKETE [4], KOGBETLIANTZ [6].
[2]) KOGBETLIANTZ [6], [7].
[3]) HARDY [5].

If r and s are positive integers, we have.

$$\Delta^r \Delta^s \varepsilon_n = \Delta^{r+s} \varepsilon_n, \tag{1.2.9}$$

and

$$\Delta^s (\delta_n \varepsilon_n) = \sum_{r=0}^{s} \binom{s}{r} \Delta^r \delta_n \Delta^{s-r} \varepsilon_{n+r}. \tag{1.2.10}$$

We also write

$$\bar{\varepsilon}_n = \Delta \varepsilon_n; \tag{1.2.11}$$

$$\begin{cases} E^{\beta,\sigma}(n,\mu) \\ \bar{E}^{\beta,\sigma}(n,\mu) \end{cases} = \sum_{\nu=0}^{n} A_{n-\nu}^{\beta-1} A_\nu^{-\sigma-1} \begin{matrix} \varepsilon_{\nu+\mu} \\ \bar{\varepsilon}_{\nu+\mu} \end{matrix}; \quad E^{\beta,\sigma}(-1,\mu) = 0, \tag{1.2.12}$$

$$\hat{E}^{\beta,\sigma}(n,\mu) = \sum_{\nu=0}^{n} A_{n-\nu}^{\beta-1} A_\nu^{-\sigma-1} (\nu+\mu)\, \varepsilon_{\nu+\mu}, \tag{1.2.13}$$

$$T_n^\varrho = \sum_{\nu=0}^{n} \frac{t_\nu^\varrho}{\nu}, \quad T_0^\varrho = 0. \tag{1.2.14}$$

Throughout this paper we use h and k to denote the integral parts of the non-negative numbers α and β respectively.

2.1. Introduction. The present paper is devoted to a generalization of the results contained in a series of three papers[4]) by PATI and the present author, concerning absolute Cesàro summability factors of infinite series.

In the said series of papers, we have extended the following theorem of BOSANQUET and CHOW in the case $\varkappa = \varrho \geqq 0$, by replacing the sequences $\{n^p\}$, $p \geqq 0$, by the monotonic non-decreasing sequence $\{\lambda_n\}$ and have also proved this more general result by means of a direct analysis which is self-contained and free from the use of apparently unconnected ancillary results of the type used by BOSANQUET and CHOW. Indeed these last mentioned authors had in essence to show the equivalence of the following theorem of theirs to a previous result of CHOW[5]).

THEOREM A[6]). If $\varkappa \geqq -1$, $\varrho \geqq 0$, $p \geqq 0$, necessary and sufficient conditions for $\sum a_n \varepsilon_n$ to be summable $|C, \varrho|$ whenever $s_n^\varkappa = O(n^p)$, as $n \to \infty$, are:

(I) a $\quad \sum n^{\varkappa-\varrho+p} |\varepsilon_n| < \infty$, $\qquad$ (I) b $\quad \sum n^{-1+p} |\varepsilon_n| < \infty$,

(II) $\quad \sum n^{\varkappa+p} |\Delta^{\varkappa+1} \varepsilon_n| < \infty$.

In this paper the author succeeds in generalizing Theorem A in the case $\varkappa \geqq 0$, $\varrho \geqq 0$, by replacing the sequence $\{n^p\}$, $p \geqq 0$, by the wider class of sequences $\{\lambda_n\}$, where $\{\lambda_n\}$ is a monotonic non-decreasing sequence.

2.2. We establish the following theorem.

THEOREM. *If $\alpha, \beta \geqq 0$, necessary and sufficient conditions for $\sum a_n \varepsilon_n$ to be summable $|C, \alpha|$ whenever $s_n^\beta = O(\lambda_n)$, as $n \to \infty$, where $\{\lambda_n\}$ is a positive monotonic non-decreasing sequence, are:*

(i) a $\quad \sum n^{\beta-\alpha} \lambda_n |\varepsilon_n| < \infty$, $\qquad$ (i) b $\quad \sum n^{-1} \lambda_n |\varepsilon_n| < \infty$,

(ii) $\quad \sum n^\beta \lambda_n |\Delta^{\beta+1} \varepsilon_n| < \infty$.

[4]) PATI and AHMAD [8], [9], [10].

[5]) CHOW [3], Theorem 2.

[6]) BOSANQUET and CHOW [2].

3.1. We require the following lemmas for the proof of our theorem.

LEMMA 1[7]). *If* $\sigma > -1$ *and* $\sigma - \delta > 0$, *then*

$$\sum_{n=\mu}^{\infty} \frac{A_{n-\mu}^{\delta}}{n A_n^{\sigma}} = \sum_{n=0}^{\infty} \frac{A_n^{\delta}}{(n+\mu) A_{n+\mu}^{\sigma}} = \frac{1}{\mu A_{\mu}^{\sigma-\delta-1}}.$$

LEMMA 2. *If* $\sigma \geqq 0$, $\delta \geqq 0$, $\{\lambda_n\}$ *be a positive monotonic non-decreasing sequence, and* $s_n^{\sigma} = O(\lambda_n)$, *then*

(i) $s_n^{\sigma+\delta} = O(\lambda_n)$

(ii) $t_n^{\sigma} = O(n\lambda_n)$, $\quad t_n^{\sigma+\delta} = O(n\lambda_n)$

and

(iii) $n a_n = O(n^{\sigma+1}\lambda_n)$.

PROOF.

(i)
$$s_n^{\sigma+\delta} = \frac{1}{A_n^{\sigma+\delta}} \sum_{\nu=0}^{n} A_{n-\nu}^{\delta-1} A_{\nu}^{\sigma} s_{\nu}^{\sigma}$$

so that

$$\begin{aligned} |s_n^{\sigma+\delta}| &\leqq \frac{1}{A_n^{\sigma+\delta}} \sum_{\nu=0}^{n} A_{n-\nu}^{\delta-1} A_{\nu}^{\sigma} |s_{\nu}^{\sigma}| \\ &\leqq \frac{K\lambda_n}{A_n^{\sigma+\delta}} \sum_{\nu=0}^{n} A_{n-\nu}^{\delta-1} A_{\nu}^{\sigma} \\ &= K\lambda_n \text{ [8]}; \end{aligned}$$

hence,

$$s_n^{\sigma+\delta} = O(\lambda_n).$$

(ii) By (1.2.1),

$$t_n^{\sigma} = n(s_n^{\sigma} - s_{n-1}^{\sigma})$$

so that

$$|t_n^{\sigma}| \leqq n(|s_n^{\sigma}| + |s_{n-1}^{\sigma}|) \leqq K n \lambda_n;$$

hence

$$t_n^{\sigma} = O(n\lambda_n).$$

The proof of the second part of (ii) is similar to the proof of (i)

(iii)
$$n a_n = \sum_{\nu=1}^{n} A_{n-\nu}^{-\sigma-1} A_{\nu}^{\sigma} t_{\mu}^{\sigma}$$

so that

$$\begin{aligned} |n a_n| &\leqq \sum_{\nu=1}^{n} |A_{n-\nu}^{-\sigma-1}| A_{\nu}^{\sigma} |t_{\mu}^{\sigma}| \\ &\leqq K(n\lambda_n) n^{\sigma} \sum_{\nu=1}^{n} |A_{n-\nu}^{-\sigma-1}| \\ &\leqq K n^{\sigma+1} \lambda_n; \end{aligned}$$

hence

$$n a_n = O(n^{\sigma+1}\lambda_n).$$

---

[7]) CHOW [3], Lemma 1.

[8]) K denotes an absolute constant, not necessarily the same at each occurrence.

LEMMA 3[9]). *Let* $\sigma \geqq 0$, $\delta \geqq 0$. *If the series* $\sum n^{-1}|t_n^{\sigma}|$ *is convergent so is the series* $\sum n^{-1}|t_n^{\sigma+\delta}|$.

LEMMA 4[10]). *If the sequences* $\{u_n\}$ *and* $\{U_n\}$ *are connected by the relation*

$$U_n = \sum_{\mu=1}^{n} D_{n,\mu}\, u_\mu \qquad (n = 1, 2, 3, \ldots),$$

a necessary condition that the series $\sum |U_n|$ *should be convergent whenever* $u_n = O(1)$ *is that*

$$\sum_{\mu=1}^{\infty} |D_{\mu,\mu}| < \infty.$$

LEMMA 5[11]). *If the sequences* $\{u_n\}$ *and* $\{U_n\}$ *are connected by the relation*

$$U_n = \sum_{\mu=1}^{\infty} D_{n,\mu}\, u_\mu. \qquad (n = 1, 2, 3, \ldots),$$

a necessary condition that the series $\sum U_n$ *should be convergent whenever* $u_n = O(1)$, *is that*

$$\sum_{\mu=1}^{\infty} \Big|\sum_{n=1}^{\infty} D_{n,\mu}\Big| < \infty.$$

LEMMA 6. *If* $\alpha \geqq 0$, $\beta \geqq 0$, *and* $\varepsilon_n = O(1)$, *then*

$$\sum = \sum_{n=0}^{\infty} \frac{\hat{E}^{\alpha,\beta+1}(n,\mu)}{(n+\mu)\, A_{n+\mu}^{\alpha}} = \Delta^{\beta+1} \varepsilon_\mu.$$

The proof is similar to that of lemma 7 of CHOW[12]). We are giving it here for completeness. We have

$$\begin{aligned}
\sum &= \sum_{n=0}^{\infty} \frac{1}{(n+\mu)\, A_{n+\mu}^{\alpha}} \sum_{\nu=0}^{n} A_{n-\nu}^{\alpha-1} A_\nu^{-\beta-2} (\nu+\mu)\, \varepsilon_{\nu+\mu} \\
&= \sum_{\nu=0}^{\infty} A_\nu^{-\beta-2} (\nu+\mu)\, \varepsilon_{\nu+\mu} \sum_{n=\nu}^{\infty} \frac{A_{n-\nu}^{\alpha-1}}{(n+\mu)\, A_{n+\mu}^{\alpha}} \\
&= \sum_{\nu=0}^{\infty} A_\nu^{-\beta-2} (\nu+\mu)\, \varepsilon_{\nu+\mu} \frac{1}{\nu+\mu} \qquad \text{(by Lemma 1)}, \\
&= \sum_{\nu=0}^{\infty} A_\nu^{-\beta-2}\, \varepsilon_{\nu+\mu} \\
&= \Delta^{\beta+1} \varepsilon_\mu,
\end{aligned}$$

the inversion of the order of summation being justified by absolute convergence.

LEMMA 7[13]). *If* $\alpha > 0$, $\beta > 0$, *then*

(i) $$E^{\alpha,\beta}(n,\mu) = \sum_{r=0}^{h} A_r^{-\beta-1} \sum_{p=0}^{r} A_p^{-r-1} \sum_{\nu=0}^{n} A_{n-\nu}^{\alpha-h-2} A_\nu^{h-r} \sum_{j=0}^{\nu} A_j^{-\beta+p-1}\, \varepsilon_{j+\mu};$$

(ii) $$\overline{E}^{\alpha,\beta}(n,\mu) = \sum_{r=0}^{h} A_r^{-\beta-1} \sum_{p=0}^{r} A_p^{-r-1} \sum_{\nu=0}^{n} A_{n-\nu}^{\alpha-h-2} A_\nu^{h-r} \sum_{=0}^{\nu} A_j^{-\beta+p-1}\, \overline{\varepsilon}_{j+\mu}.$$

[9]) KOGBETLIANTZ [6].

[10]) CHOW [3], Lemma 6.

[11]) CHOW [3], Lemma 5.

[12]) CHOW [3].

[13]) (i) has been proved in CHOW [3], Lemma 12, the proof of (ii) is exactly similar.

LEMMA 8[14]). *Let* $\alpha \geqq 0$. *If* $\sum \lambda_n |\varepsilon_n| < \infty$, *and* $\sum n^{\alpha} \lambda_n |\Delta^{\alpha+1} \varepsilon_n| < \infty$; *then*

(i) $$\sum_{\mu=1}^{\infty} A_{\mu}^{\alpha} \lambda_{\mu} \sum_{n=0}^{\infty} \frac{|E^{\alpha,\alpha}(n,\mu)|}{(n+\mu) A_{n+\mu}^{\alpha}} < \infty,$$

(ii) $$\sum_{\mu=1}^{\infty} A_{\mu}^{\alpha+1} \lambda_{\mu} \sum_{n=1}^{\infty} \frac{|\bar{E}^{\alpha,\alpha}(n,\mu)|}{(n+\mu) A_{n+\mu}^{\alpha}} < \infty,$$

(iii) $$\sum_{\mu=1}^{\infty} A_{\mu}^{\alpha+1} \lambda_{\mu} \sum_{n=1}^{\infty} \frac{|E^{\alpha-1,\alpha}(n,\mu+1)|}{(n+\mu) A_{n+\mu}^{\alpha}} < \infty.$$

LEMMA 9. *If* $p \geqq 0$, $\{\lambda_n\}$ *is a positive monotonic non-decreasing sequence, and* $\sum n^{p} \lambda_n |\Delta^{p+1} \varepsilon_n| < \infty$, *then* $\sum n^{q} \lambda_n |\Delta^{q+1} \varepsilon_n| < \infty$, *for every* q *such that* $0 \leqq q \leqq p$.

This is proved by an easy adaptation of the proof of Lemma 2 [1].

3.2. Proof of the Theorem.

Let $\tau_n^0 = \tau_n = n\, a_n\, \varepsilon_n$, and let τ_n^{α} ($\alpha > 0$) denote the n-th Cesàro mean of order α of the sequence $\{\tau_n\}$. Then by definition and by (1.2.1), the series $\sum a_n \varepsilon_n$ is summable $|C, \alpha|$ if and only if

$$\sum n^{-1} |\tau_n^{\alpha}| < \infty.$$

NECESSITY. Since, by (1.2.13),

$$\begin{aligned}
\hat{E}^{\alpha,\beta+1}(n,\mu) &= \sum_{\nu=0}^{n} A_{n-\nu}^{\alpha-1} A_{\nu}^{-\beta-2} (\nu+\mu)\, \varepsilon_{\nu+\mu} \\
&= \sum_{\nu=0}^{n-1} \Delta_{\nu} [A_{n-\nu}^{\alpha-1} (\nu+\mu)\, \varepsilon_{\nu+\mu}]\, A_{\nu}^{-\beta-1} + (n+\mu)\, \varepsilon_{n+\mu}\, A_n^{-\beta-1} \\
&= \sum_{\nu=0}^{n} \Delta_{\nu} [A_{n-\nu}^{\alpha-1} (\nu+\mu)\, \varepsilon_{\nu+\mu}]\, A_{\nu}^{-\beta-1}.
\end{aligned}$$

We have

$$\begin{aligned}
A_n^{\alpha} \tau_n^{\alpha} &= \sum_{\nu=1}^{n} A_{n-\nu}^{\alpha-1} \nu\, a_{\nu}\, \varepsilon_{\nu} \\
&= \sum_{\nu=1}^{n-1} \Delta_{\nu} [A_{n-\nu}^{\alpha-1} \nu\, \varepsilon_{\nu}]\, s_{\nu} + n\, \varepsilon_n\, s_n \\
&= \sum_{\nu=1}^{n} \Delta_{\nu} [A_{n-\nu}^{\alpha-1} \nu\, \varepsilon_{\nu}]\, s_{\nu} \\
&= \sum_{\nu=1}^{n} \Delta_{\nu} [A_{n-\nu}^{\alpha-1} \nu\, \varepsilon_{\nu}] \sum_{\mu=1}^{\nu} A_{\nu-\mu}^{-\beta-1} A_{\mu}^{\beta} s_{\mu}^{\beta} \\
&= \sum_{\mu=1}^{n} A_{\mu}^{\beta} s_{\mu}^{\beta} \sum_{\nu=0}^{n-\mu} \Delta_{\nu} [A_{n-\mu-\nu}^{\alpha-1} (\nu+\mu)\, \varepsilon_{\nu+\mu}]\, A_{\nu}^{-\beta-1}.
\end{aligned}$$

Hence

(3.2.1) $$U_n = \frac{\tau_n^{\alpha}}{n} = \sum_{\mu=1}^{n} \left(\frac{s_{\mu}^{\beta}}{\lambda_n}\right) \frac{\lambda_{\mu} A_{\mu}^{\beta}}{n A_n^{\alpha}} \hat{E}^{\alpha,\beta+1}(n-\mu,\mu).$$

[14]) PATI and AHMAD [8], [9], [10].

CONDITION (i)b IS NECESSARY. For if $\sum a_n \varepsilon_n$ is summable $|C|$ whenever $s_n^\beta = O(\lambda_n)$, as $n \to \infty$, then, by Lemma 2, it is summable $|C|$ whenever $n a_n = O(\lambda_n)$, as $n \to \infty$.

Suppose that $\sum n^{-1} \lambda_n |\varepsilon_n| = \infty$. Let

$$a_n = n^{-1} \lambda_n \operatorname{sgn} \varepsilon_n,$$

then

$$n a_n = \lambda_n \operatorname{sgn} \varepsilon_n = O(\lambda_n).$$

But

$$\begin{aligned} \sum a_n \varepsilon_n &= \sum n^{-1} \lambda_n \varepsilon_n \operatorname{sgn} \varepsilon_n \\ &= \sum n^{-1} \lambda_n |\varepsilon_n| = \infty, \end{aligned}$$

so that $\sum a_n \varepsilon_n$ is not summable $|C|$. This is a contradiction.

CONDITION (i)a IS NECESSARY. In order that $\sum n^{-1} |\tau_n^\alpha| < \infty$, whenever $s_n^\beta = O(\lambda_n)$, as $n \to \infty$, by (3.2.1) and lemma 4, it is necessary that

$$\sum_{\mu=1}^{\infty} \frac{A_\mu^\beta \lambda_\mu}{\mu A_\mu^\alpha} |\hat{E}^{\alpha,\beta+1}(\mu - \mu, \mu)| < \infty,$$

which is equivalent to

$$\sum_{\mu=1}^{\infty} \mu^{\beta-\alpha} \lambda_\mu |\varepsilon_\mu| < \infty.$$

CONDITION (ii) IS NECESSARY. For (i)a and (i)b both imply $\varepsilon_n = O(1)$. Then, in order that $\sum n^{-1} |\tau_n^\alpha| < \infty$, whenever $s_n^\beta = O(\lambda_n)$, as $n \to \infty$, by (3.2.1) and Lemma 5, it is necessary that

$$\sum_{\mu=1}^{\infty} \left| \sum_{n=\mu}^{\infty} \frac{A_\mu^\beta \lambda_\mu}{n A_n^\alpha} \hat{E}^{\alpha,\beta+1}(n - \mu, \mu) \right| < \infty,$$

i.e. that

$$\sum_{\mu=1}^{\infty} \mu^\beta \lambda_\mu \left| \sum_{n=0}^{\infty} \frac{\hat{E}^{\alpha,\beta+1}(n, \mu)}{(n + \mu) A_{n+\mu}^\alpha} \right| < \infty,$$

which, by Lemma 6, is equivalent to

$$\sum_{\mu=1}^{\infty} \mu^\beta \lambda_\mu |\Delta^{\beta+1} \varepsilon_\mu| < \infty.$$

This completes the proof of the necessity part of the theorem.

SUFFICIENCY.

Since, by hypothesis and by (1.2.1),

$$T_n^\beta = O(\lambda_n), \quad \text{as} \quad n \to \infty.$$

We have

$$\begin{aligned} A_n^\alpha \tau_n^\alpha &= \sum_{\nu=1}^{n} A_{n-\nu}^{\alpha-1} \nu a_\nu \varepsilon_\nu \\ &= \sum_{\nu=1}^{n} A_{n-\nu}^{\alpha-1} \varepsilon_\nu \sum_{\mu=1}^{\nu} A_{\nu-\mu}^{-\beta-1} A_\mu^\beta t_\mu^\beta \\ &= \sum_{\mu=1}^{n} A_\mu^\beta t_\mu^\beta \sum_{\nu=0}^{n-\mu} A_{n-\mu-\nu}^{\alpha-1} A_\nu^{-\beta-1} \varepsilon_{\nu+\mu} \\ &= \sum_{\mu=1}^{n} \{(\mu A_\mu^\beta) E^{\alpha,\beta}(n - \mu, \mu)\} \left(\frac{t_\mu^\beta}{\mu}\right). \end{aligned}$$

Hence, by Abel's transformation,

$$A_n^\alpha \tau_n^\alpha = \sum_{\mu=1}^{n-1} \Delta_\mu \{(\mu A_\mu^\beta) E^{\alpha,\beta}(n-\mu,\mu)\} T_\mu^\beta + n A_n^\beta E^{\alpha,\beta}(0,n) T_n^\beta$$

$$= \sum_{\mu=1}^{n} \Delta_\mu \{(\mu A_\mu^\beta) E^{\alpha,\beta}(n-\mu,\mu)\} T_\mu^\beta + (n+1) A_{n+1}^\beta E^{\alpha,\beta}(-1, n+1) T_n^\beta$$

$$= \sum_{\mu=1}^{n} \Delta_\mu \{(\mu A_\mu^\beta) E^{\alpha,\beta}(n-\mu,\mu)\} T_\mu^\beta$$

$$= \sum_{\mu=1}^{n} \Delta_\mu (\mu A_\mu^\beta) E^{\alpha,\beta}(n-\mu,\mu) T_\mu^\beta + \sum_{\mu=1}^{n} (\mu+1) A_{\mu+1}^\beta \Delta_\mu E^{\alpha,\beta}(n-\mu,\mu) T_\mu^\beta.$$

Now

$$\Delta_\mu E^{\alpha,\beta}(n-\mu,\mu)$$

$$= \sum_{\nu=0}^{n-\mu-1} A_\nu^{-\beta-1} \Delta_\mu (A_{n-\mu-\nu}^{\alpha-1} \varepsilon_{\nu+\mu}) + A_{n-\mu}^{-\beta-1} \varepsilon_n$$

$$= \sum_{\nu=0}^{n-\mu} A_\nu^{-\beta-1} A_{n-\mu-\nu}^{\alpha-1} \Delta \varepsilon_{\nu+\mu} + \sum_{\nu=0}^{n-\mu} A_\nu^{-\beta-1} A_{n-\mu-\nu}^{\alpha-2} \varepsilon_{\nu+\mu+1} -$$

$$- A_{n-\mu}^{-\beta-1} \Delta \varepsilon_n - A_{n-\mu}^{-\beta-1} \varepsilon_{n+1} + A_{n-\mu}^{-\beta-1} \varepsilon_n$$

$$= \bar{E}^{\alpha,\beta}(n-\mu,\mu) + E^{\alpha-1,\beta}(n-\mu,\mu+1).$$

Hence

$$A_n^\alpha \tau_n^\alpha = \sum_{\mu=1}^{n} \Delta_\mu (\mu A_\mu^\beta) E^{\alpha,\beta}(n-\mu,\mu) T_\mu^\beta + \sum_{\mu=1}^{n} (\mu+1) A_{\mu+1}^\beta \bar{E}^{\alpha,\beta}(n-\mu,\mu) T_\mu^\beta +$$

$$+ \sum_{\mu=1}^{n} (\mu+1) A_{\mu+1}^\beta E^{\alpha-1,\beta}(n-\mu,\mu+1) T_\mu^\beta.$$

Therefore

$$\sum_{n=1}^{\infty} n^{-1} |\tau_n^\alpha| \leq \sum_{n=1}^{\infty} (n A_n^\alpha)^{-1} \sum_{\mu=1}^{n} |\Delta_\mu (\mu A_\mu^\beta)| \, |E^{\alpha,\beta}(n-\mu,\mu)| \, |T_\mu^\beta| +$$

$$+ \sum_{n=1}^{\infty} (n A_n^\alpha)^{-1} \sum_{\mu=1}^{n} (\mu+1) A_{\mu+1}^\beta |\bar{E}^{\alpha,\beta}(n-\mu,\mu)| \, |T_\mu^\beta| +$$

$$+ \sum_{n=1}^{\infty} (n A_n^\alpha)^{-1} \sum_{\mu=1}^{n} (\mu+1) A_{\mu+1}^\beta |E^{\alpha-1,\beta}(n-\mu,\mu+1)| \, |T_\mu^\beta|.$$

Thus, it is sufficient to prove that

$$\text{(3.2.2)} \qquad \begin{cases} \sum_1 = \sum_{\mu=1}^{\infty} A_\mu^\beta \lambda_\mu \sum_{n=0}^{\infty} \frac{|E^{\alpha,\beta}(n,\mu)|}{(n+\mu) A_{n+\mu}^\alpha} < \infty; \\ \sum_2 = \sum_{\mu=1}^{\infty} A_\mu^{\beta+1} \lambda_\mu \sum_{n=0}^{\infty} \frac{|\bar{E}^{\alpha,\beta}(n,\mu)|}{(n+\mu) A_{n+\mu}^\alpha} < \infty; \\ \sum_3 = \sum_{\mu=1}^{\infty} A_\mu^{\beta+1} \lambda_\mu \sum_{n=0}^{\infty} \frac{|E^{\alpha-1,\beta}(n,\mu+1)|}{(n+\mu) A_{n+\mu}^\alpha} < \infty. \end{cases}$$

Proof of (3.2.2.).

CASE I. *When* $0=\alpha<\beta$. We have

$$\begin{aligned}\sum_1 &\leqq \sum_{\mu=1}^{\infty} A_\mu^{\beta}\lambda_\mu \sum_{n=0}^{\infty}|A_n^{-\beta-1}|\,\frac{\varepsilon_{n+\mu}}{n+\mu}\Big| \\ &\leqq \sum_{n=1}^{\infty}\Big|\frac{\varepsilon_n}{n}\Big|\lambda_n\sum_{\mu=1}^{n}|A_{n-\mu}^{-\beta-1}|\,A_\mu^{\beta} \\ &\leqq K\sum_{n=1}^{\infty} n^{\beta-1}\lambda_n|\varepsilon_n|<\infty,\end{aligned}$$

by hypothesis.

Next,

$$\begin{aligned}\sum_2 &\leqq \sum_{\mu=1}^{\infty} A_\mu^{\beta+1}\lambda_\mu \sum_{n=0}^{\infty}|A_n^{-\beta-1}|\,\frac{|\bar{\varepsilon}_{n+\mu}|}{n+\mu} \\ &\leqq \sum_{n=1}^{\infty}\frac{|\bar{\varepsilon}_n|}{n}\lambda_n\sum_{\mu=1}^{n}|A_{n-\mu}^{-\beta-1}|\,A_\mu^{\beta+1} \\ &\leqq K\sum_{n=1}^{\infty} n^{\beta}\lambda_n|\Delta\varepsilon_n| \\ &\leqq K\sum_{n=1}^{\infty} n^{\beta}\lambda_n|\varepsilon_n|+K\sum_{n=1}^{\infty} n^{\beta}\lambda_n|\varepsilon_{n+1}| \\ &<\infty,\end{aligned}$$

by hypothesis.

Lastly,

$$\begin{aligned}\sum_3 &\leqq \sum_{\mu=1}^{\infty} A_\mu^{\beta+1}\lambda_\mu \sum_{n=0}^{\infty}|A_n^{-\beta-1}|\,\frac{|\varepsilon_{n+\mu+1}|}{n+\mu}-\sum_{\mu=1}^{\infty} A_\mu^{\beta+1}\lambda_\mu\sum_{n=0}^{\infty}|A_{n-1}^{-\beta-1}|\,\frac{|\varepsilon_{n+\mu}|}{n+\mu} \\ &\leqq \sum_{n=1}^{\infty}\frac{|\varepsilon_{n+1}|}{n}\lambda_n\sum_{\mu=1}^{n}|A_{n-\mu}^{-\beta-1}|\,A_\mu^{\beta+1}-\sum_{n=1}^{\infty}\frac{|\varepsilon_{n+1}|}{n+1}\lambda_n\sum_{\mu=1}^{n}|A_{n-\mu}^{-\beta-1}|\,A_\mu^{\beta+1} \\ &\leqq K\sum_{n=1}^{\infty} n^{\beta}\lambda_n|\varepsilon_{n+1}|+K\sum_{n=1}^{\infty} n^{\beta}\lambda_n|\varepsilon_{n+1}| \\ &<\infty,\end{aligned}$$

by hypothesis.

CASE II. *When* $0<\alpha\leqq\beta$. We have

$$\begin{aligned}\sum_1 &= \sum_{\mu=1}^{\infty} A_\mu^{\beta}\lambda_\mu\Big(\sum_{n=0}^{k}+\sum_{n=k+1}^{\infty}\Big)\frac{|E^{\alpha,\beta}(n,\mu)|}{(n+\mu)\,A_{n+\mu}^{\alpha}} \\ &= \sum_{1,1}+\sum_{1,2}, \text{ say,} \\ \sum_2 &= \sum_{\mu=1}^{\infty} A_\mu^{\beta+1}\lambda_\mu\Big(\sum_{n=0}^{k}+\sum_{n=k+1}^{\infty}\Big)\frac{|\bar{E}^{\alpha,\beta}(n,\mu)|}{(n+\mu)\,A_{n+\mu}^{\alpha}} \\ &= \sum_{2,1}+\sum_{2,2}, \text{ say,}\end{aligned}$$

and

$$\sum_3 = \sum_{\mu=1}^{\infty} A_\mu^{\beta+1} \lambda_\mu \Big(\sum_{n=0}^{k+1} + \sum_{n=k+2}^{\infty} \Big) \frac{|E^{\alpha-1,\beta}(n, \mu+1)|}{(n+\mu) A_{n+\mu}^{\alpha}}$$

$$= \sum_{3,1} + \sum_{3,2}, \text{ say.}$$

Hence, it is sufficient for our purpose to show that

(3.2.3) $$\sum_{s,1} < \infty, \qquad (s = 1, 2, 3);$$

(3.2.4) $$\sum_{s,2} < \infty, \qquad (s = 1, 2, 3).$$

Proof of (3.2.3.)

We have

$$\sum_{1,1} \leqq \sum_{\mu=1}^{\infty} A_\mu^{\beta} \lambda_\mu \sum_{n=0}^{k} \frac{1}{(n+\mu) A_{n+\mu}^{\alpha}} \sum_{\nu=0}^{n} A_{n-\nu}^{\alpha-1} |A_\nu^{-\beta-1}| \, |\varepsilon_{\nu+\mu}|$$

$$= \sum_{\mu=1}^{\infty} A_\mu^{\beta} \lambda_\mu \sum_{n=0}^{k} |A_\nu^{-\beta-1}| \, |\varepsilon_{\nu+\mu}| \sum_{n=\nu}^{k} \frac{A_{n-\nu}^{\alpha-1}}{(n+\mu) A_{n+\mu}^{\alpha}}$$

$$\leqq \sum_{\mu=1}^{\infty} A_\mu^{\beta} \lambda_\mu \sum_{\nu=0}^{k} \frac{|A_\nu^{-\beta-1}| \, |\varepsilon_{\nu+\mu}|}{(\nu+\mu) A_{\nu+\mu}^{\alpha}} \sum_{n=0}^{k-\nu} A_n^{\alpha-1}$$

$$\leqq \sum_{\mu=1}^{\infty} \frac{A_\mu^{\beta} \lambda_\mu}{\mu A_\mu^{\alpha}} \sum_{\nu=0}^{k} |A_\nu^{-\beta-1}| \, |\varepsilon_{\nu+\mu}| \, A_{k-\nu}^{\alpha}$$

$$= \sum_{\nu=0}^{k} O\Big(\sum_{\mu=1}^{\infty} \mu^{\beta-\alpha-1} \lambda_\mu |\varepsilon_{\nu+\mu}| \Big)$$

$$< \infty,$$

by hypothesis.

Next, we have

$$\sum_{2,1} \leqq \sum_{\mu=1}^{\infty} A_\mu^{\beta+1} \lambda_\mu \sum_{n=0}^{k} \frac{1}{(n+\mu) A_{n+\mu}^{\alpha}} \sum_{\nu=0}^{n} A_{n-\nu}^{\alpha-1} |A_\nu^{-\beta-1}| \, |\bar{\varepsilon}_{\nu+\mu}|$$

$$\leqq \sum_{\mu=1}^{\infty} \frac{A_\mu^{\beta+1} \lambda_\mu}{\mu A_\mu^{\alpha}} \sum_{\nu=0}^{k} |A_\nu^{-\beta-1}| \, |\bar{\varepsilon}_{\nu+\mu}| \sum_{n=0}^{k-\nu} A_n^{\alpha-1}$$

$$= \sum_{\nu=0}^{k} O\Big(\sum_{\mu=1}^{\infty} \mu^{\beta-\alpha} \lambda_\mu |\Delta \varepsilon_{\nu+\mu}| \Big)$$

$$= \sum_{\nu=0}^{k} O\Big(\sum_{\mu=1}^{\infty} \mu^{\beta-\alpha} \lambda_\mu |\varepsilon_{\nu+\mu}| \Big) +$$

$$+ \sum_{\nu=0}^{k} O\Big(\sum_{\mu=1}^{\infty} \mu^{\beta-\alpha} \lambda_\mu |\varepsilon_{\nu+\mu+1}| \Big)$$

$$< \infty,$$

by hypothesis.

Lastly, we have

$$\begin{aligned}\sum_{3,1} &\leqq \sum_{\mu=1}^{\infty} A_\mu^{\beta+1}\lambda_\mu \sum_{n=0}^{k+1} \frac{1}{(n+\mu)\,A_{n+\mu}^{\alpha}} \sum_{\nu=0}^{n} |A_{n-\nu}^{\alpha-2}|\,|A_\nu^{-\beta-1}|\,|\varepsilon_{\nu+\mu+1}| \\ &= \sum_{\mu=1}^{\infty} A_\mu^{\beta+1}\lambda_\mu \sum_{\nu=0}^{k+1} |A_\nu^{-\beta-1}|\,|\varepsilon_{\nu+\mu+1}| \sum_{n=\nu}^{k+1} \frac{|A_{n-\nu}^{\alpha-2}|}{(n+\mu)\,A_{n+\mu}^{\alpha}} \\ &\leqq \sum_{\mu=1}^{\infty} \frac{A_\mu^{\beta+1}\lambda_\mu}{\mu\,A_\mu^{\alpha}} \sum_{\nu=0}^{k+1} |A_\nu^{-\beta-1}|\,|\varepsilon_{\nu+\mu+1}| \sum_{n=0}^{k+1-\nu} |A_n^{\alpha-2}| \\ &\leqq K \sum_{\mu=1}^{\infty} \frac{A_\mu^{\beta+1}\lambda_\mu}{\mu\,A_\mu^{\alpha}} \sum_{\nu=0}^{k+1} |A_\nu^{-\beta-1}|\,|\varepsilon_{\nu+\mu+1}| \\ &= \sum_{\nu=0}^{k+1} O\Big(\sum_{\mu=1}^{\infty} \mu^{\beta-\alpha}\lambda_\mu |\varepsilon_{\nu+\mu+1}|\Big) \\ &< \infty,\end{aligned}$$

by hypothesis. This completes the proof of (3.2.3).

Proof of (3.2.4).

CASE (II_1). *When* $0<\alpha<\beta$.

By Lemma 7(i), we have

$$(3.2.5)\,a \qquad |E^{\alpha,\beta}(n,\mu)| = O\Big(\sum_{r=0}^{k}\sum_{p=0}^{r} |Q^{\alpha,\beta}(r,p;n,\mu)|\Big)$$

where

$$Q^{\alpha,\beta}(r,p;n,\mu) = \sum_{\nu=0}^{n} A_{n-\nu}^{\alpha-k-2}\,A_\nu^{k-r} \sum_{j=0}^{\nu} A_j^{-\beta+p-1}\,\varepsilon_{j+\mu}.$$

We write

$$(3.2.6)\,a \qquad Q^{\alpha,\beta}(r,p;n,\mu) = \sum_{s=1}^{3} Q_s^{\alpha,\beta}(n,\mu),$$

where

$$(3.2.7)\,a \qquad Q_1^{\alpha,\beta}(n,\mu) = A_n^{\alpha-r-1}\,\Delta^{\beta-p}\,\varepsilon_\mu,$$

$$(3.2.8)\,a \qquad Q_2^{\alpha,\beta}(n,\mu) = -A_n^{\alpha-r-1} \sum_{j=n+1}^{\infty} A_j^{-\beta+p-1}\,\varepsilon_{j+\mu},$$

$$(3.2.9)\,a \qquad Q_3^{\alpha,\beta}(n,\mu) = -\sum_{\nu=0}^{n} A_{n-\nu}^{\alpha-k-2}\,A_\nu^{k-r} \sum_{j=\nu+1}^{n} A_j^{-\beta+p-1}\,\varepsilon_{j+\mu}.$$

Hence, in order to prove that $\sum_{1,2}<\infty$, by (3.2.5)a and (3.2.6)a, it is sufficient to show that, for $0\leqq p\leqq r\leqq k$,

$$(3.2.10)\,a \qquad \sum_{1,2,s} = \sum_{\mu=1}^{\infty} A_\mu^{\beta}\lambda_\mu \sum_{n=k+1}^{\infty} \frac{|Q_s^{\alpha,\beta}(n,\mu)|}{(n+\mu)\,A_{n+\mu}^{\alpha}} < \infty, \qquad (s=1,2,3).$$

Proof of (3.2.10) a.

We can easily see that, for $n>k$, $Q_2^{\alpha,\beta}(n,\mu)\equiv 0$ when β is an integer, and $Q_3^{\alpha,\beta}(n,\mu)\equiv 0$ when α and β are both integers.

Since $0 \leqq p \leqq r \leqq h < \alpha < \beta$. By (3.2.7) a, we have

$$\begin{aligned}
\sum_{1,2,1} &= \sum_{\mu=1}^{\infty} A_\mu^\beta \lambda_\mu \sum_{n=h+1}^{\infty} \frac{A_n^{\alpha-r-1} |\Delta^{\beta-p} \varepsilon_\mu|}{(n+\mu) A_{n+\mu}^\alpha} \\
&\leqq \sum_{\mu=1}^{\infty} A_\mu^\beta \lambda_\mu |\Delta^{\beta-p} \varepsilon_\mu| \sum_{n=0}^{\infty} \frac{A_n^{\alpha-r-1}}{(n+\mu) A_{n+\mu}^\alpha} \\
&= \sum_{\mu=1}^{\infty} A_\mu^\beta \lambda_\mu |\Delta^{\beta-p} \varepsilon_\mu| \frac{1}{\mu A_\mu^r} \qquad \text{(by Lemma 1)} \\
&\leqq K \sum_{\mu=1}^{\infty} \mu^{\beta-p-1} \lambda_\mu |\Delta^{\beta-p} \varepsilon_\mu| \\
&< \infty,
\end{aligned}$$

by hypothesis and Lemma 9.

Next, by (3.2.8) a, we have

$$\begin{aligned}
\sum_{1,2,2} &\leqq \sum_{\mu=1}^{\infty} A_\mu^\beta \lambda_\mu \sum_{n=0}^{\infty} \frac{A_n^{\alpha-r-1}}{(n+\mu) A_{n+\mu}^\alpha} \sum_{j=n+1}^{\infty} |A_j^{-\beta+p-1}| \, |\varepsilon_{j+\mu}| \\
&= \sum_{\mu=1}^{\infty} A_\mu^\beta \lambda_\mu \sum_{j=1}^{\infty} |A_j^{-\beta+p-1}| \, |\varepsilon_{j+\mu}| \sum_{n=0}^{j-1} \frac{A_n^{\alpha-r-1}}{(n+\mu) A_{n+\mu}^\alpha} \\
&\leqq K \sum_{\mu=1}^{\infty} \mu^{\beta-\alpha-1} \lambda_\mu \sum_{j=1}^{\infty} |A_j^{-\beta+\alpha-1}| \, |\varepsilon_{j+\mu}| \\
&\leqq K \sum_{j=2}^{\infty} \lambda_j |\varepsilon_j| \sum_{\mu=1}^{j-1} \mu^{\beta-\alpha-1} |A_{j-\mu}^{-\beta+\alpha-1}| \\
&\leqq K \sum_{j=2}^{\infty} \lambda_j |\varepsilon_j| \left| \sum_{\mu=0}^{j} A_\mu^{\beta-\alpha-1} A_{j-\mu}^{-\beta+\alpha-1} \right| + \\
&\quad + K \sum_{j=2}^{\infty} j^{\beta-\alpha-1} \lambda_j |\varepsilon_j| + K \sum_{j=2}^{\infty} j^{-\beta+\alpha-1} \lambda_j |\varepsilon_j| \\
&< \infty,
\end{aligned}$$

by hypothesis.

Lastly, by (3.2.9) a, we have

$$\begin{aligned}
\sum_{1,2,3} &\leqq \sum_{\mu=1}^{\infty} A_\mu^\beta \lambda_\mu \sum_{n=1}^{\infty} \frac{1}{(n+\mu) A_{n+\mu}^\alpha} \sum_{\nu=0}^{n} |A_{n-\nu}^{\alpha-h-2}| \, A_\nu^{h-r} \sum_{j=\nu+1}^{n} |A_j^{-\beta+p-1}| \, |\varepsilon_{j+\mu}| \\
&= \sum_{\mu=1}^{\infty} A_\mu^\beta \lambda_\mu \sum_{n=1}^{\infty} \frac{1}{(n+\mu) A_{n+\mu}^\alpha} \sum_{j=1}^{n} |A_j^{-\beta+p-1}| \, |\varepsilon_{j+\mu}| \sum_{\nu=0}^{j-1} |A_{n-\nu}^{\alpha-h-2}| \, A_\nu^{h-r} \\
&\leqq \sum_{\mu=1}^{\infty} A_\mu^\beta \lambda_\mu \sum_{j=1}^{\infty} |A_j^{-\beta+p-1}| \, |\varepsilon_{j+\mu}| \sum_{\nu=0}^{j-1} A_\nu^{h-r} \sum_{n=j}^{\infty} \frac{|A_{n-\nu}^{\alpha-h-2}|}{(n+\mu) A_{n+\mu}^\alpha} \\
&\leqq \sum_{\mu=1}^{\infty} A_\mu^\beta \lambda_\mu \sum_{j=1}^{\infty} \frac{|A_j^{-\beta+p-1}| \, |\varepsilon_{j+\mu}|}{(j+\mu) A_{j+\mu}^\alpha} \sum_{\nu=0}^{j-1} A_{j-\nu}^{\alpha-h-1} A_\nu^{h-r} \\
&\leqq K \sum_{\mu=1}^{\infty} A_\mu^\beta \lambda_\mu \sum_{j=1}^{\infty} |A_j^{-\beta+\alpha-1}| \, |\varepsilon_{j+\mu}| \frac{1}{(j+\mu) A_{j+\mu}^\alpha}
\end{aligned}$$

$$\leqq K \sum_{j=2}^{\infty} \frac{\lambda_j |\varepsilon_j|}{j A_j^{\alpha}} \sum_{\mu=1}^{j-1} A_{\mu}^{\beta} |A_{j-\mu}^{-\beta+\alpha-1}|$$

$$\leqq K \sum_{j=2}^{\infty} j^{\beta-\alpha-1} \lambda_j |\varepsilon_j|$$

$$< \infty,$$

by hypothesis. This completes the proof of (3.2.10) a.

Again, by Lemma 7 (ii), we have

(3.2.5) b $$|\bar{E}^{\alpha,\beta}(n,\mu)| = O\Big(\sum_{r=0}^{h} \sum_{p=0}^{r} |\bar{Q}^{\alpha,\beta}(r,p;n,\mu)|\Big)$$

where

$$\bar{Q}^{\alpha,\beta}(r,p;n,\mu) = \sum_{\nu=0}^{n} A_{n-\nu}^{\alpha-h-2} A_{\nu}^{h-r} \sum_{j=0}^{\nu} A_j^{-\beta+p-1} \bar{\varepsilon}_{j+\mu}.$$

We write

(3.2.6) b $$\bar{Q}^{\alpha,\beta}(r,p;n,\mu) = \sum_{s=0}^{3} \bar{Q}_s^{\alpha,\beta}(n,\mu),$$

where

(3.2.7) b $$\bar{Q}_1^{\alpha,\beta}(n,\mu) = A_n^{\alpha-r-1} \Delta^{\beta-p} \bar{\varepsilon}_{\mu},$$

(3.2.8) b $$\bar{Q}_2^{\alpha,\beta}(n,\mu) = -A_n^{\alpha-r-1} \sum_{j=n+1}^{\infty} A_j^{-\beta+p-1} \bar{\varepsilon}_{j+\mu},$$

(3.2.9) b $$\bar{Q}_3^{\alpha,\beta}(n,\mu) = -\sum_{\nu=0}^{n} A_{n-\nu}^{\alpha-h-2} A_{\nu}^{h-r} \sum_{j=\nu+1}^{n} A_j^{-\beta+p-1} \bar{\varepsilon}_{j+\mu}.$$

Hence, in order to prove that $\sum_{2,2} < \infty$, by (3.2.5) b and (3.2.6) b, it is sufficient to show that, for $0 \leqq p \leqq r \leqq h$,

(3.2.10) b $$\sum_{2,2,s} = \sum_{\mu=0}^{\infty} A_{\mu}^{\beta+1} \lambda_{\mu} \sum_{n=h+1}^{\infty} \frac{|\bar{Q}_s^{\alpha,\beta}(n,\mu)|}{(n+\mu) A_{n+\mu}^{\alpha}} < \infty, \qquad (s = 1, 2, 3).$$

Proof of (3.2.10) b.

We can easily see that, for $n > k$, $\bar{Q}_2^{\alpha,\beta}(n,\mu) \equiv 0$ when β is an integer, and $\bar{Q}_3^{\alpha,\beta}(n,\mu) \equiv 0$ when α and β are both integers. Since $0 \leqq p \leqq r \leqq h < \alpha < \beta$. First, by (3.2.7) a, we have

$$\sum_{2,2,1} = \sum_{\mu=1}^{\infty} A_{\mu}^{\beta+1} \lambda_{\mu} \sum_{n=k+1}^{\infty} \frac{A_n^{\alpha-r-1} |\Delta^{\beta-p} \bar{\varepsilon}_{\mu}|}{(n+\mu) A_{n+\mu}^{\alpha}}$$

$$\leqq \sum_{\mu=1}^{\infty} A_{\mu}^{\beta+1} \lambda_{\mu} |\Delta^{\beta-p} \bar{\varepsilon}_{\mu}| \sum_{n=0}^{\infty} \frac{A_n^{\alpha-r-1}}{(n+\mu) A_{n+\mu}^{\alpha}}$$

$$= \sum_{\mu=1}^{\infty} A_{\mu}^{\beta+1} \lambda_{\mu} |\Delta^{\beta-p+1} \varepsilon_{\mu}| \frac{1}{\mu A_{\mu}^{r}} \qquad \text{(by Lemma 1)}$$

$$\leqq K \sum_{\mu=1}^{\infty} \mu^{\beta-p} \lambda_{\mu} |\Delta^{\beta-p+1} \varepsilon_{\mu}|$$

$$< \infty,$$

by hypothesis and Lemma 9.

Next, by (3.2.8) b, we have

$$\begin{aligned}
\sum_{2,2,2} &\leqq \sum_{\mu=1}^{\infty} A_{\mu}^{\beta+1}\,\lambda_{\mu} \sum_{n=0}^{\infty} \frac{A_n^{\alpha-r-1}}{(n+\mu)\,A_{n+\mu}^{\alpha}} \sum_{j=n+1}^{\infty} |A_j^{-\beta+r-1}|\,|\bar{\varepsilon}_{j+\mu}| \\
&= \sum_{\mu=1}^{\infty} A_{\mu}^{\beta+1}\,\lambda_{\mu} \sum_{j=1}^{\infty} |A_j^{-\beta+r-1}|\,|\bar{\varepsilon}_{j+\mu}| \sum_{n=0}^{j-1} \frac{A_n^{\alpha-r-1}}{(n+\mu)\,A_{n+\mu}^{\alpha}} \\
&\leqq K \sum_{\mu=1}^{\infty} \mu^{\beta-\alpha}\,\lambda_{\mu} \sum_{j=1}^{\infty} |A_j^{-\beta+\alpha-1}|\,|\bar{\varepsilon}_{j+\mu}| \\
&\leqq K \sum_{j=2}^{\infty} \lambda_j |\bar{\varepsilon}_j| \sum_{\mu=1}^{j-1} \mu^{\beta-\alpha}\,|A_{j-\mu}^{-\beta+\alpha-1}| \\
&\leqq K \sum_{j=2}^{\infty} j^{\beta-\alpha}\,\lambda_j |\Delta\,\varepsilon_j| \\
&\leqq K \sum_{j=2}^{\infty} j^{\beta-\alpha}\,\lambda_j |\varepsilon_j| + K \sum_{j=2}^{\infty} j^{\beta-\alpha}\,\lambda_j |\varepsilon_{j+1}| \\
&< \infty,
\end{aligned}$$

by hypothesis.

Lastly, by (3.2.9) b, we have

$$\begin{aligned}
\sum_{2,2,3} &\leqq \sum_{\mu=1}^{\infty} A_{\mu}^{\beta+1}\,\lambda_{\mu} \sum_{n=1}^{\infty} \frac{1}{(n+\mu)\,A_{n+\mu}^{\alpha}} \sum_{\nu=0}^{n} |A_{n-\nu}^{\alpha-k-2}|\,A_{\nu}^{k-r} \sum_{j=\nu+1}^{n} |A_j^{-\beta+r-1}|\,|\bar{\varepsilon}_{j+\mu}| \\
&= \sum_{\mu=1}^{\infty} A_{\mu}^{\beta+1}\,\lambda_{\mu} \sum_{n=1}^{\infty} \frac{1}{(n+\mu)\,A_{n+\mu}^{\alpha}} \sum_{j=1}^{n} |A_j^{-\beta+r-1}|\,|\bar{\varepsilon}_{j+\mu}| \sum_{\nu=0}^{j-1} |A_{n-\nu}^{\alpha-k-2}|\,A_{\nu}^{k-r} \\
&\leqq \sum_{\mu=1}^{\infty} A_{\mu}^{\beta+1}\,\lambda_{\mu} \sum_{j=1}^{\infty} |A_j^{-\beta+r-1}|\,|\bar{\varepsilon}_{j+\mu}| \sum_{\nu=0}^{j-1} A_{\nu}^{k-r} \sum_{n=j}^{\infty} \frac{|A_{n-\nu}^{\alpha-k-2}|}{(n+\mu)\,A_{n+\mu}^{\alpha}} \\
&\leqq \sum_{\mu=1}^{\infty} A_{\mu}^{\beta+1}\,\lambda_{\mu} \sum_{j=1}^{\infty} \frac{|A_j^{-\beta+r-1}|\,|\bar{\varepsilon}_{j+\mu}|}{(j+\mu)\,A_{j+\mu}^{\alpha}} \sum_{\nu=0}^{j-1} A_{j-\nu}^{\alpha-k-1}\,A_{\nu}^{k-r} \\
&\leqq K \sum_{\mu=1}^{\infty} A_{\mu}^{\beta+1}\,\lambda_{\mu} \sum_{j=1}^{\infty} |A_j^{-\beta+\alpha-1}|\,|\bar{\varepsilon}_{j+\mu}|\,\frac{1}{(j+\mu)\,A_{j+\mu}^{\alpha}} \\
&\leqq K \sum_{j=2}^{\infty} \lambda_j \frac{|\bar{\varepsilon}_j|}{j\,A_j^{\alpha}} \sum_{\mu=1}^{j-1} A_{\mu}^{\beta+1} |A_{j-\mu}^{-\beta+\alpha-1}| \\
&\leqq K \sum_{j=2}^{\infty} j^{\beta-\alpha}\,\lambda_j |\Delta\,\varepsilon_j| \\
&\leqq K \sum_{j=2}^{\infty} j^{\beta-\alpha}\,\lambda_j |\varepsilon_j| + K \sum_{j=2}^{\infty} j^{\beta-\alpha}\,\lambda_j |\varepsilon_{j+1}| \\
&< \infty,
\end{aligned}$$

by hypothesis. This completes the proof of (3.2.10) b.

Finally, by Lemma 7(i), we have

$$(3.2.5)\,c \qquad |E_n^{\alpha-1,\beta}(n,\mu+1)| = O\Big(\sum_{r=0}^{k}\sum_{p=0}^{r}|Q^{\alpha-1,\beta}(r,p;n,\mu+1)|\Big)$$

where

$$Q^{\alpha-1,\beta}(r,p;n,\mu+1) = \sum_{\nu=0}^{n} A_{n-\nu}^{\alpha-k-3} A_\nu^{k-r} \sum_{j=0}^{\nu} A_j^{-\beta+p-1}\,\varepsilon_{j+\mu+1}.$$

We write

$$(3.2.6)\,c \qquad Q^{\alpha-1,\beta}(r,p;n,\mu+1) = \sum_{s=1}^{3} Q_s^{\alpha-1,\beta}(n,\mu+1),$$

where

$$(3.2.7)\,c \qquad Q_1^{\alpha-1,\beta}(n,\mu+1) = A_n^{\alpha-r-2}\,\Delta^{\beta-p}\,\varepsilon_{\mu+1},$$

$$(3.2.8)\,c \qquad Q_2^{\alpha-1,\beta}(n,\mu+1) = -\,A_n^{\alpha-r-2}\sum_{j=n+1}^{\infty} A_j^{-\beta+p-1}\,\varepsilon_{j+\mu+1},$$

$$(3.2.9)\,c \qquad Q_3^{\alpha-1,\beta}(n,\mu+1) = -\sum_{\nu=0}^{n} A_{n-\nu}^{\alpha-k-3} A_\nu^{k-r} \sum_{j=\nu+1}^{n} A_j^{-\beta+p-1}\,\varepsilon_{j+\mu+1}.$$

Hence, in order to prove that $\sum_{3,2}<\infty$, by (3.2.5)c and (3.2.6)c, it is sufficient to show that, for $0\leqq p\leqq r\leqq h$,

$$(3.2.10)\,c \qquad \sum_{3,2,s} = \sum_{\mu=1}^{\infty} A_\mu^{\beta+1}\,\lambda_\mu \sum_{n=k+2}^{\infty} \frac{|Q_s^{\alpha-1,\beta}(n,\mu+1)|}{(n+\mu)\,A_{n+\mu}^{\alpha}} < \infty, \qquad (s=1,2,3).$$

Proof of (3.2.10)c.

We can easily see that, for $n>k+1$, $Q_2^{\alpha-1,\beta}(n,\mu+1)\equiv 0$ when β is an integer, and $Q_3^{\alpha-1,\beta}(n,\mu+1)\equiv 0$ when α and β are both integers.

Since $0\leqq p\leqq r\leqq h<\alpha<\beta$. First, by (3.2.7)c, we have

$$\begin{aligned}
\sum_{3,2,1} &= \sum_{\mu=1}^{\infty} A_\mu^{\beta+1}\,\lambda_\mu \sum_{n=k+2}^{\infty} \frac{|A_n^{\alpha-r-2}||\Delta^{\beta-p}\,\varepsilon_{\mu+1}|}{(n+\mu)\,A_{n+\mu}^{\alpha}} \\
&\leqq \sum_{\mu=1}^{\infty} A_\mu^{\beta+1}\,\lambda_\mu |\Delta^{\beta-p}\,\varepsilon_{\mu+1}| \sum_{n=0}^{\infty} \frac{|A_n^{\alpha-r-2}|}{(n+\mu)\,A_{n+\mu}^{\alpha}} \\
&\leqq K\sum_{\mu=1}^{\infty} \mu^{\beta-\alpha}\,\lambda_\mu |\Delta^{\beta-p}\,\varepsilon_{\mu+1}| \\
&\leqq K\sum_{\mu=1}^{\infty} \mu^{\beta-\alpha}\,\lambda_\mu \sum_{\nu=0}^{\infty} |A_\nu^{-\beta+p-1}|\,|\varepsilon_{\nu+\mu+1}| \\
&\leqq K\sum_{\nu=1}^{\infty} \lambda_\nu\,|\varepsilon_{\nu+1}| \sum_{\mu=1}^{\nu} |A_{\nu-\mu}^{-\beta+p-1}|\,\mu^{\beta-\alpha} \\
&\leqq K\sum_{\nu=1}^{\infty} \nu^{\beta-\alpha}\,\lambda_\nu\,|\varepsilon_{\nu+1}| \\
&< \infty,
\end{aligned}$$

by hypothesis.

Next, by (3.2.8)c, we have

$$\begin{aligned}\sum_{3,2,2} &\leqq \sum_{\mu=1}^{\infty} A_\mu^{\beta+1}\,\lambda_\mu \sum_{n=0}^{\infty} \frac{|A_n^{\alpha-r-2}|}{(n+\mu)\,A_{n+\mu}^{\alpha}} \sum_{j=n+1}^{\infty} |A_j^{-\beta+p-1}|\,|\varepsilon_{j+\mu+1}| \\ &= \sum_{\mu=1}^{\infty} A_\mu^{\beta+1}\,\lambda_\mu \sum_{j=1}^{\infty} |A_j^{-\beta+p-1}|\,|\varepsilon_{j+\mu+1}| \sum_{n=0}^{j-1} \frac{|A_n^{\alpha-r-2}|}{(n+\mu)\,A_{n+\mu}^{\alpha}} \\ &\leqq K \sum_{\mu=1}^{\infty} \mu^{\beta-\alpha}\,\lambda_\mu \sum_{j=1}^{\infty} |A_j^{-\beta+p-1}|\,|\varepsilon_{j+\mu+1}| \\ &\leqq K \sum_{j=2}^{\infty} \lambda_j\,|\varepsilon_{+1}| \sum_{\mu=1}^{j-1} \mu^{\beta-\alpha} |A_{j-\mu}^{-\beta+p-1}| \\ &\leqq K \sum_{j=2}^{\infty} j^{\beta-\alpha}\,\lambda_j\,|\varepsilon_{j+1}| \\ &< \infty,\end{aligned}$$

by hypothesis.

Lastly, by (3.2.9)c, we have

$$\begin{aligned}\sum_{3,2,3} &\leqq \sum_{\mu=1}^{\infty} A_\mu^{\beta+1}\,\lambda_\mu \sum_{n=0}^{\infty} \frac{1}{(n+\mu)\,A_{n+\mu}^{\alpha}} \sum_{\nu=0}^{n} |A_{n-\nu}^{\alpha-k-3}|\,A_\nu^{k-r} \sum_{j=\nu+1}^{n} |A_j^{-\beta+p-1}|\,|\varepsilon_{j+\mu+1}| \\ &= \sum_{\mu=1}^{\infty} A_\mu^{\beta+1}\,\lambda_\mu \sum_{n=1}^{\infty} \frac{1}{(n+\mu)\,A_{n+\mu}^{\alpha}} \sum_{j=1}^{n} |A_j^{-\beta+p-1}|\,|\varepsilon_{j+\mu+1}| \sum_{\nu=0}^{j-1} |A_{n-\nu}^{\alpha-k-3}|\,A_\nu^{k-r} \\ &\leqq \sum_{\mu=1}^{\infty} A_\mu^{\beta+1}\,\lambda_\mu \sum_{j=1}^{\infty} |A_j^{-\beta+p-1}|\,|\varepsilon_{j+\mu+1}| \sum_{\nu=0}^{j-1} A_\nu^{k-r} \sum_{n=j}^{\infty} \frac{|A_{n-\nu}^{\alpha-k-3}|}{(n+\mu)\,A_{n+\mu}^{\alpha}} \\ &\leqq \sum_{\mu=1}^{\infty} A_\mu^{\beta+1}\,\lambda_\mu \sum_{j=1}^{\infty} \frac{|A_j^{-\beta+p-1}|\,|\varepsilon_{j+\mu+1}|}{(j+\mu)\,A_{j+\mu}^{\alpha}} \sum_{\nu=0}^{j-1} |A_{j-\nu}^{\alpha-k-2}|\,A_\nu^{k-r} \\ &\leqq K \sum_{\mu=1}^{\infty} A_\mu^{\beta+1}\,\lambda_\mu \sum_{j=1}^{\infty} |A_j^{-\beta+k-1}|\,|\varepsilon_{j+\mu+1}|\,\frac{1}{(j+\mu)\,A_{j+\mu}^{\alpha}} \\ &\leqq K \sum_{j=2}^{\infty} \lambda_j \frac{|\varepsilon_{j+1}|}{j\,A_j^{\alpha}} \sum_{\mu=1}^{j-1} A_\mu^{\beta+1}\,|A_{j-\mu}^{-\beta+k-1}| \\ &\leqq K \sum_{j=2}^{\infty} j^{\beta-\alpha}\,\lambda_j\,|\varepsilon_{j+1}| \\ &< \infty,\end{aligned}$$

by hypothesis. This completes the proof of (3.2.10)c and thus, completes the proof of (3.2.4) for the case $0<\alpha<\beta$.

CASE (II_2). *When* $\alpha=\beta\geqq 0$ [15].

We see that, by Lemma 8, the inequalities (3.2.4) hold in this case also.

[15]) For this case, the theorem has recently been proved by PATI and AHMAD, in the series of papers, [8], [9], [10].

CASE III. *When* $\alpha > \beta \geqq 0$. Then, by what we have proved, $\sum a_n \varepsilon_n$ is summable $|C, \beta|$. By Lemmas 2 and 3, it is also summable $|C, \alpha|$.

This completes the proof of the sufficiency part of the theorem, and also the proof of the theorem.

I should like to express my warmest thanks to Dr. T. PATI for his kind encouragement and valuable guidance.

References

[1] BOSANQUET, L. S.: Note on the Bohr-Hardy theorem. J. London Math. Soc. **17**, 166–173 (1942).

[2] —, and H. C. CHOW: Some remarks on convergence and summability factors. J. London Math. Soc. **32**, 73–82 (1957).

[3] CHOW, H. C.: Note on convergence and summability factors. J. London Math. Soc. **29**, 459–476 (1954).

[4] FEKETE, M.: Zur Theorie der divergenten Reihen. Math. es termesz értesito (Budapest) **29**, 719–726 (1911).

[5] HARDY, G. H.: Divergent Series. Oxford 1949.

[6] KOGBETLIANTZ, E.: Sur les séries absolument sommables par la méthode des moyennes arithmétiques. Bull. Sci. Math. (2) **49**, 234–256 (1925).

[7] — Sommation des séries et intégrales divergentes par les moyennes arithmétiques et typiques. Mémorial Sci. Math. No. 51 (1931).

[8] PATI, T., and Z. U. AHMAD: On the absolute summability factors of infinite series I. Tôhoku Math. J. **12**, 222–232 (1960).

[9] — — On the absolute summability factors of infinite series II. Indian J. Math. **2**, 29–39 (1960).

[10] — — On the absolute summability factors of infinite series III. Indian J. Math. **2**, 73–87 (1960).

University of Allahabad, Allahabad, India

(Eingegangen am 1. September 1960)

ANDRÉ, J.
Math. Zeitschr. 76, 311—333 (1961)

Über Parallelstrukturen *)

Teil IV: T-Strukturen[1])

Von

JOHANNES ANDRE

Einleitung

Als T-Struktur definierten wir in III, §1 eine affine Parallelstruktur (a.P.), deren Translationen alle zentral sind und in der es in jeder Richtung mindestens eine nichtidentische Translation gibt. Die zentralen Translationsstrukturen (z.T.; vgl. II, §2) sind Beispiele von T-Strukturen.

Da jede T-Struktur auch t-Struktur ist, gelten alle in III gebrachten Sätze auch für T-Strukturen. So ist insbesondere jede T-Struktur translationstreu über einer (bis auf Isomorphie) eindeutig bestimmten z.T., ihrem Stamm. Wir wollen jetzt folgendes Problem genauer untersuchen: *Gegeben sei eine z.T. t, gesucht ist eine vollständige Übersicht über alle T-Strukturen, die translationstreue Erweiterungen von t sind.* In III sind einige translationstreue Erweiterungen von t gebildet worden; diese, wie etwa die freien translationstreuen Erweiterungsebenen (III, §§3 bis 6) sind aber keine T-Strukturen. Nunmehr werden wir ein Verfahren angeben, durch das wir *alle* translationstreuen Erweiterungen von t aufstellen können, soweit sie T-Strukturen sind. Wir müssen hierzu aus einer (endlichen oder unendlichen) Anzahl paarweise isomorpher disjunkt liegender z.T. die T-Strukturen zusammenbauen; das wird in §2 durchgeführt. Wesentliches Hilfsmittel hierfür ist eine „Matrix" von Abbildungen der Geradenmenge des Stammes auf sich oder die leere Menge, die „Matrix der kanonischen Abbildungen". Ehe wir diese Konstruktion durchführen, gehen wir in §1 den umgekehrten Weg und fragen, in welcher Beziehung die von einer vorgegebenen T-Struktur induzierten z.T. (III, §2, Ziffer 2) untereinander stehen. Wesentliches Hilfsmittel sind auch hier die kanonischen Abbildungen. Die Untersuchung der T-Strukturen wird damit zurückgeführt auf die der z.T. und der kanonischen Abbildungen. Die Ergebnisse werden dazu benutzt, um in §3 allgemeine Sätze über T-Strukturen, insbesondere über deren Kollineationsgruppen, herzuleiten.

Die letzten beiden Paragraphen bringen spezielle Beispiele und Anwendungen der vorher gebrachten Theorie, auch auf affine Ebenen. Es wird gezeigt, daß es T-Strukturen mit genau zwei Transitivitätsgebieten hinsichtlich der Translationsgruppe nicht geben kann. Dagegen existieren T-Strukturen

*) Teil einer von der Naturwissenschaftlich-Philosophischen Fakultät der Technischen Hochschule Braunschweig angenommenen Habilitationsschrift.

[1]) Teil I, Math. Z. **76**, 85—102, Teil II, Math. Z. **76**, 155—163, Teil III, Math. Z. **76**, 240—256 (1961); im folgenden hier mit I, II bzw. III zitiert.

mit genau drei Transitivitätsgebieten; Beispiele hierfür werden in §4 angegeben, insbesondere werden noch einige nähere Angaben gemacht über Anzahl der Punkte und Geraden der kleinsten bisher bekannten T-Struktur, die keine Translationsstruktur ist. Schließlich wird noch gezeigt, daß es endliche T-Ebenen (das sind T-Strukturen, die zugleich affine Ebenen sind) mit genau zwei Transitivitätsgebieten hinsichtlich der von den Translationen und Strekkungen erzeugten Gruppe nicht geben kann [vgl. auch ANDRÉ (1958)]; das ist ein Spezialfall eines allgemeinen Satzes über die Nichtexistenz bestimmter endlicher T-Strukturen (Satz 5.2). Es ist anzunehmen, daß die hier gebrachten Beweismethoden auch noch weitere Aussagen über T-Ebenen liefern können.

Die Kenntnis der §§3 bis 6 von III wird für diesen Teil nicht vorausgesetzt.

§1. Die kanonischen Abbildungen

1. Sei $\mathfrak{p}=(P, \mathfrak{G})$ eine T-Struktur, d.h. eine t-Struktur mit nichtidentischen Translationen in jeder Richtung. Durch ihre Translationsgruppe T wird P in $|I|$ Transitivitätsgebiete T_i $(i \in I)$ zerlegt. Die durch $\mathfrak{p}$ auf T_i induzierten z.T. (III, §2, Ziffer 2) seien $\mathfrak{t}_i=(T_i, \mathfrak{L}_i)$. Nach III, Satz 2.3, sind sie zum Stamm $\mathfrak{t}=(T, \mathfrak{L})$ isomorph[2]). Ferner sei ι_i ein (für das folgende festgehaltener) kanonischer Isomorphismus von $\mathfrak{t}$ auf $\mathfrak{t}_i$ [vgl. III, (2.4)]. Die Spur $G \cap T_i$ einer Geraden $G \in \mathfrak{G}$ besteht nach III, Satz 2.2, aus mindestens zwei Punkten.

Sei $L \in \mathfrak{L}$, ihr Bild $L_i=L^{\iota_i} \in \mathfrak{L}_i$ ist Spur einer eindeutig bestimmten Geraden $L^i \in \mathfrak{G}$, d.h. es gilt $L_i=L^i \cap T_i$ (vgl. III, §2, Ziffer 2). Wir definieren jetzt auf $\mathfrak{L}$ eine Abbildung $\varkappa_{ij}$ $(i, j \in I)$ durch

$$L^{\varkappa_{ij}}=(L^i \cap T_j)^{\iota_j^{-1}} \tag{1.1}$$

und nennen $\varkappa_{ij}$ eine *kanonische Abbildung*.

2. Ehe wir die kennzeichnenden Eigenschaften der kanonischen Abbildungen aufstellen, müssen wir noch einen Begriff einführen, der — wie wir später sehen werden — im gewissen Sinne als eine Verallgemeinerung der Partition aufgefaßt werden kann.

DEFINITION. Eine *Zerlegung* $\mathfrak{Z}$ einer z.T. $\mathfrak{t}=(T, \mathfrak{L})$ ist eine Menge paarweise windschiefer Geraden von $\mathfrak{t}$, die ganz T überdecken[3]).

Durch jeden Punkt von $\mathfrak{t}$ geht also genau eine Gerade von $\mathfrak{Z}$, und zwei verschiedene Geraden aus $\mathfrak{Z}$ sind nicht zueinander parallel. Mit $\mathfrak{Z}$ ist auch $\mathfrak{Z}^\gamma$ eine Zerlegung (wobei γ eine Kollineation von $\mathfrak{t}$ ist). Ferner gilt der folgende nützliche

SATZ 1.1. *Jede Zerlegung enthält mindestens drei Geraden.*

BEWEIS. Zunächst folgt aus Axiom (a 3) für a.P., daß eine Zerlegung $\mathfrak{Z}$ von $\mathfrak{t}$ (falls eine solche überhaupt existiert) mindestens zwei Geraden L und L' enthält. Sei $p \in L'$; ein von p verschiedener Punkt $q \in \langle L \| p \rangle$ (zur Definition

[2]) Die Punktmenge T von $\mathfrak{t}$ stimmt natürlich mit T überein (vgl. die Definition des Stammes in III, §2); die Abänderung erfolgt lediglich aus Gründen der einheitlichen Bezeichnung, vgl. hierzu die in Ziffer 4 der Einleitung von I gemachten Bemerkungen.

[3]) *Beispiele* von Zerlegungen werden in §4 gebracht.

von $\langle L \| p \rangle$ s. I, §2, Ziffer 2) existiert nach (t 2) und gehört weder zu L noch zu L'. Die durch q gehende Gerade $L'' \in \mathfrak{Z}$ ist eine gewünschte dritte Gerade.

3. Nunmehr stellen wir einige Eigenschaften der kanonischen Abbildungen zusammen (von denen im nächsten Paragraphen gezeigt wird, daß sie für diese Abbildungen kennzeichnend sind).

SATZ 1.2. *Die durch* (1.1) *definierten kanonischen Abbildungen* $\varkappa_{ij}$ $(i, j \in I)$ *einer z. T.* $t = (T, \mathfrak{L})$ *genügen den folgenden Bedingungen:*

(I) $L^{\varkappa_{ij}}$ *ist entweder leer oder eine zu* L *parallele Gerade aus* $\mathfrak{L}$.

(II) *Für alle* $L \in \mathfrak{L}$ *und* $i \in I$ *gilt* $L^{\varkappa_{ii}} = L$ (d. h. $\varkappa_{ii} = \underline{1}$).

(III) *Aus* $L^{\varkappa_{ij}} \neq \emptyset$ *folgt* $L^{\varkappa_{ij}\varkappa_{jk}} = L^{\varkappa_{ik}}$ $(i, j, k \in I)$.

(IV) *Jedes* $\varkappa_{ij}$ *ist mit jeder Translation* τ *von* t *vertauschbar:* $\varkappa_{ij}\tau = \tau\varkappa_{ij}$.

(V) *Ist* $\mathfrak{P}$ *die Menge aller durch einen festen Punkt gehenden Geraden von* t, *so stellt* $\mathfrak{P}^{\varkappa_{ij}}$ *für* $i \neq j$ *eine Zerlegung von* t *dar*[4]).

(VI) *Zu jeder auf* I *regulären Permutation*[5]) $\varrho \neq \underline{1}$ *und jeder Streckung* σ *von* t *gibt es* $i, j \in I$ *mit* $\varkappa_{i\varrho, j\varrho} \neq \sigma^{-1}\varkappa_{ij}\sigma$.

BEWEIS. Sei T die Translationsgruppe von $\mathfrak{p} = (P, \mathfrak{G})$; wir können und wollen sie mit der von t identifizieren, denn $\mathfrak{p}$ ist als translationstreu über t vorausgesetzt. Wir beweisen jetzt nacheinander die Richtigkeit der Bedingungen (I) bis (VI).

(I): Wegen (1.1) gilt dann genau $L^{\varkappa_{ij}} \neq \emptyset$, wenn $L^i \cap T_j \neq \emptyset$ $(i, j \in I)$ ist; dies wollen wir jetzt voraussetzen. Wenn $\tau \in \mathsf{T}(L^i)$ genommen wird, also $L^{i\tau} = L^i$ und dann ebenso $L^\tau = L$ gilt, so folgt $L^{\varkappa_{ij}\tau} = (L^i \cap T_j)^{\iota_j^{-1}\tau} = (L^i \cap T_j)^{\tau\iota_j^{-1}} = (L^i \cap T_j)^{\iota_j^{-1}} = L^{\varkappa_{ij}}$, somit $L^{\varkappa_{ij}} \| L$, da ja t eine z. T. ist (vgl. II, §2). Die benutzte Vertauschbarkeit von ι_i^{-1} mit τ folgt aus $\iota_i\tau = \tau\iota_i$ und diese Beziehung aus der Definition des kanonischen Isomorphismus ι_i [vgl. III, (2.4)], denn für jedes $t \in T$ gilt $t = o^{\bar{t}}$ [wobei $\bar{t}$ die t zugeordnete Translation vermöge des zwischen T und T bestehenden Isomorphismus ist, vgl. II, §1, Ziffer 2, insbesondere (1.1)], also ($o^{\varkappa_{ij}} = o_i$ gesetzt) $t^{\varkappa_{ij}} = o_i^{\bar{t}}$, $t^{\varkappa_{ij}\tau} = o_i^{\bar{t}\tau} = o_i^{\overline{t^\tau}} = t^{\tau\varkappa_{ij}}$.

(II) folgt aus $L^{\varkappa_{ii}} = (L^i \cap T_i)^{\iota_i^{-1}} = L_i^{\iota_i^{-1}} = L$.

(III): Wir setzen $L^{\varkappa_{ij}} \neq \emptyset$, also $L^i \cap T_j \neq \emptyset$ voraus, erhalten $L^i = \big((L^i \cap T_j)^{\iota_j^{-1}}\big)^j = L^{\varkappa_{ij}j}$ und hieraus schließlich $L^{\varkappa_{ij}\varkappa_{jk}} = (L^{\varkappa_{ij}j} \cap T_k)^{\iota_k^{-1}} = (L^i \cap T_k)^{\iota_k^{-1}} = L^{\varkappa_{ik}}$.

(IV) schließt man ähnlich wie (I): $L^{\varkappa_{ij}\tau} = (L^i \cap T_j)^{\iota_j^{-1}\tau} = (L^i \cap T_j)^{\tau\iota_j^{-1}} = (L^{i\tau} \cap T_j)^{\iota_j^{-1}} = (L^{\tau i} \cap T_j)^{\iota_j^{-1}} = L^{\tau\varkappa_{ij}}$.

(V): Wir müssen zeigen, daß $\mathfrak{P}^{\varkappa_{ij}}$ für $i \neq j$ die Eigenschaften einer Zerlegung hat.

---

[4]) Hierbei ist $\mathfrak{P}^{\varkappa_{ij}}$ als die Menge derjenigen $L^{\varkappa_{ij}}$ $(L \in \mathfrak{P})$ definiert, *die nicht leer sind.* Diese Definition, die von der üblichen [nach der auch $\emptyset$ zu $\mathfrak{P}^{\varkappa_{ij}}$ gehören würde, wie man aus (I) leicht herleiten kann] etwas abweicht, erweist sich im folgenden als zweckmäßig.

[5]) Eine Permutation ϱ auf I heißt *regulär*, wenn jede Potenz ϱ^n von ϱ alle oder kein Element von I festläßt [ZASSENHAUS (1937), S. 35]. Ist I endlich, so ist ϱ genau dann regulär, wenn alle ihre Zyklen die gleiche Länge haben.

(a) Seien $p, q \in T$ und $\mathfrak{P}$ die Menge aller durch p gehenden Geraden aus $\mathfrak{L}$; ferner sei $p^{\iota_i} = p_i$ und $q^{\iota_j} = q_j$ gesetzt; wegen $i \neq j$ ist $p_i \neq q_j$. Die Spur $\langle p_i, q_j \rangle_{\mathfrak{p}} \cap T_i$ von $\langle p_i, q_j \rangle_{\mathfrak{p}}$ auf T_i enthält mindestens zwei Punkte, denn $\mathfrak{p}$ ist T-Struktur. Somit ist $L = (\langle p_i, q_j \rangle_{\mathfrak{p}} \cap T_i)^{\iota_i^{-1}}$ eine Gerade aus $\mathfrak{L}$. Wegen $p \in L$ ist $L \in \mathfrak{P}$, und aus $\langle p_i, q_j \rangle_{\mathfrak{p}} = L^{\iota_i}$ folgt, daß $L^{\varkappa_{ij}} = (\langle p_i, q_j \rangle_{\mathfrak{p}} \cap T_j)^{\iota_j^{-1}}$ eine durch q gehende Gerade aus $\mathfrak{P}^{\varkappa_{ij}}$ ist.

(b) Sei nun umgekehrt $q \in L' \in \mathfrak{P}^{\varkappa_{ij}}$; dann gibt es ein $L'' \in \mathfrak{P}$ mit $L' = L''^{\varkappa_{ij}} = (L''^{\iota_i} \cap T_j)^{\iota_j^{-1}}$. Wegen $p \in L''$ und $q \in L'$ sind $p_i, q_j \in L''^{\iota_i}$, also gilt $L''^{\iota_i} = \langle p_i, q_j \rangle_{\mathfrak{p}} = L^{\iota_i}$, $L'' = L$, $L' = L^{\varkappa_{ij}}$; damit ist gezeigt, daß nur eine Gerade aus $\mathfrak{P}^{\varkappa_{ij}}$ durch q geht.

(c) Seien $L^{\varkappa_{ij}}$ und $L'^{\varkappa_{ij}}$ zwei verschiedene Geraden aus $\mathfrak{P}^{\varkappa_{ij}}$. Wegen $p \in L$, L' und $L \neq L'$ ist $L \nparallel L'$. Aus (II), (III) und (I) folgt $L = L^{\varkappa_{ii}} = L^{\varkappa_{ij}\varkappa_{ji}} \parallel L^{\varkappa_{ij}}$, ebenso $L' \parallel L^{\varkappa_{ij}}$, hieraus $L^{\varkappa_{ij}} \nparallel L'^{\varkappa_{ij}}$.

Damit sind für $\mathfrak{P}^{\varkappa_{ij}}$ alle Eigenschaften einer Zerlegung von $\mathfrak{k}$ nachgewiesen.

(VI): Nehmen wir an, (VI) sei nicht erfüllt. Dann gibt es eine auf I reguläre Permutation $\varrho \neq \underline{1}$ und eine Streckung σ von $\mathfrak{k}$ derart, daß $\varkappa_{i\varrho, j\varrho} = \sigma^{-1} \varkappa_{ij} \sigma$ für alle $i, j \in I$ gilt. Wir zeigen zunächst, daß die durch

$$t_j^{\varrho^*} = t^{\iota_j \varrho^*} = t^{\sigma \iota_{\bar{j}}} \qquad (t \in T,\ j \in I,\ \bar{j} = j^{\varrho} \text{ gesetzt}) \tag{1.2}$$

auf P definierte Abbildung ϱ^* in $\mathfrak{p}$ eine Translation darstellt. Aus der Definition (1.2) folgt wegen $\varrho \neq \underline{1}$, also $j^{\varrho} \neq j$ für jedes $j \in I$, sofort, daß ϱ^* fixpunktfrei ist. Zu jedem $G \in \mathfrak{G}$ gibt es ein $i \in I$ und ein $L \in \mathfrak{L}$ mit $G = L^{\iota_i} = \bigcup_{j \in I} (L^{\iota_i} \cap T_j) = \bigcup_{j \in I} L^{\varkappa_{ij} \iota_j}$, denn $\mathfrak{p}$ ist T-Struktur. Wegen $\iota_j \varrho^* = \sigma \iota_{j\varrho}$ [vgl. (1.2)] erhalten wir:

$$G^{\varrho^*} = \bigcup_{j \in I} L^{\varkappa_{ij} \iota_j \varrho^*} = \bigcup_{j \in I} L^{\varkappa_{ij} \sigma \iota_{j\varrho}} = \bigcup_{j \in I} L^{\sigma \sigma^{-1} \varkappa_{ij} \sigma \iota_{j\varrho}}$$

$$= \bigcup_{j \in I} L^{\sigma \varkappa_{i\varrho, j\varrho} \iota_{j\varrho}} = \bigcup_{j \in I} (L^{\sigma \iota_{i\varrho}} \cap T_{j\varrho}) = L^{\sigma \iota_{i\varrho}} \parallel L^{\iota_i} = G;$$

die Parallelität zwischen G und G^{ϱ^*} folgt daraus, daß deren Spuren in T_i bzw. $T_{i\varrho}$, nämlich L^{ι_i} und $L^{\sigma \iota_{i\varrho}}$, zueinander parallel sind (als Bilder der in $\mathfrak{k}$ parallelen Geraden L und L^{σ}). Somit ist ϱ^* als Translation von $\mathfrak{p}$ erkannt. Andererseits ist $T_i^{\varrho^*} = T_{i\varrho} \neq T_i$ wegen $\varrho \neq \underline{1}$, das ist aber ein Widerspruch zu unserer Annahme, daß T_i ein Transitivitätsgebiet hinsichtlich der Gruppe T *aller* Translationen von $\mathfrak{p}$ sei.

4. Wir bringen jetzt einige Folgerungen aus Satz 1.2, zu deren Beweis wir nur die Bedingungen (I) bis (VI) benutzen, dagegen nicht mehr unmittelbar auf die Definition (1.1) der kanonischen Abbildungen zurückgreifen[6]).

(a) *Aus* $L^{\varkappa_{ij}} \neq \emptyset$ *und* $L \parallel L'$ *folgt* $L'^{\varkappa_{ij}} \neq \emptyset$.

Aus $L \parallel L'$ folgt nämlich die Existenz eines $\tau \in \mathsf{T}$ mit $L' = L^{\tau}$, denn $\mathfrak{k}$ ist Translationsstruktur. Nach (IV) ist also $L'^{\varkappa_{ij}} = L^{\tau \varkappa_{ij}} = L^{\varkappa_{ij} \tau} \neq \emptyset$. Ebenso ergibt sich:

[6]) Dies ist wichtig für das im folgenden Paragraphen durchzuführende Umkehrproblem: Aus gegebenen $\varkappa_{ij}$ mit den Eigenschaften (I) bis (VI) eine T-Struktur zu konstruieren.

(b) *Aus* $L^{\varkappa_{ij}} = L$ *und* $L \| L'$ *folgt* $L'^{\varkappa_{ij}} = L'$.

(c) *Aus* $L^{\varkappa_{ij}} \neq \emptyset$ *und* $L^{\varkappa_{jk}} \neq \emptyset$ *folgt* $L^{\varkappa_{ik}} \neq \emptyset$.

Aus (I) folgt nämlich $L \| L^{\varkappa_{ij}}$ und aus (III) und (a) nunmehr $L^{\varkappa_{ik}} = L^{\varkappa_{ij}\varkappa_{jk}} \neq \emptyset$.

Aus (II) und (III) ergibt sich sofort:

(d) *Aus* $L^{\varkappa_{ij}} \neq \emptyset$ *folgt* $L^{\varkappa_{ij}\varkappa_{ji}} = L$.

(e) *Aus* $L^{\varkappa_{ij}} \neq \emptyset$ *folgt* $L^{\varkappa_{ji}} \neq \emptyset$.

Wäre nämlich $L^{\varkappa_{ji}} = \emptyset$, so wegen $L^{\varkappa_{ij}} \| L$ nach (a) auch $L^{\varkappa_{ij}\varkappa_{ji}} = \emptyset$, was im Widerspruch zu (d) steht.

Wir bezeichnen mit $\mathfrak{L}(i, j)$ die Menge aller $L \in \mathfrak{L}$, für die $L^{\varkappa_{ij}} \neq \emptyset$ ist. Dann können wir (e) auch so formulieren:

$$\mathfrak{L}(i, j) = \mathfrak{L}(j, i), \tag{1.3}$$

und für (c) können wir schreiben:

$$\mathfrak{L}(i, j) \cap \mathfrak{L}(j, k) \subseteq \mathfrak{L}(i, k). \tag{1.4}$$

Aus (II) folgt sofort $\mathfrak{L}(i, i) = \mathfrak{L}$.

(f) *Nach* (d) *und* (e) *ist* $\varkappa_{ij}$ *eine eineindeutige Abbildung von* $\mathfrak{L}(i, j)$ *auf sich, und es gilt dort*

$$\varkappa_{ij}^{-1} = \varkappa_{ji}. \tag{1.5}$$

Die nächste Folgerung kehrt in gewisser Weise (b) um:

(g) *Aus* $i \neq j$, $L^{\varkappa_{ij}} = L$ *und* $L'^{\varkappa_{ij}} = L'$ *folgt* $L \| L'$.

Beweis. Sei $p \in L$. Wäre $L \nparallel L'$, so $L'' = \langle L' \| p \rangle \neq L$. Aus (b) folgt andererseits $L''^{\varkappa_{ij}} = L''$, also hätten wir $\{p\} = L \cap L'' = L^{\varkappa_{ij}} \cap L''^{\varkappa_{ij}}$, was für $i \neq j$ im Widerspruch zu (V) steht. — Aus Satz 1.1, (V) und (g) folgt sofort

(h) *Zu jedem* $\varkappa_{ij}$ *mit* $i \neq j$ *gibt es* $L, L' \in \mathfrak{L}(i, j)$ *mit* $L \nparallel L'$ *derart, daß* $L^{\varkappa_{ij}} \neq L$ *und* $L'^{\varkappa_{ij}} \neq L'$ *ist.*

(i) *Zu jedem* $\varkappa_{ij}$ *und* $L \in \mathfrak{L}(i, j)$ *gibt es eine Translation* $\tau_{ij,L}$ *mit*

$$L^{\varkappa_{ij}} = L^{\tau_{ij,L}}. \tag{1.6}$$

Das folgt sofort aus (I) und daraus, daß $\mathfrak{k}$ Translationsstruktur ist.

(j) *Aus* $L \| L'$ *folgt* $\tau_{ij,L} \equiv \tau_{ij,L'}$ $(\mathsf{T}(L))$.

Aus der Voraussetzung folgt nämlich die Existenz eines $\tau \in \mathsf{T}$ mit $L' = L^{\tau}$. Weil T abelsch ist, folgt $L'^{\tau_{ij,L'}} = L'^{\varkappa_{ij}} = L^{\tau\varkappa_{ij}} = L^{\varkappa_{ij}\tau} = L^{\tau_{ij,L}\tau} = L^{\tau\tau_{ij,L}} = L'^{\tau_{ij,L}}$. Entsprechend schließt man $L''^{\tau_{ij,L'}} = L''^{\tau_{ij,L}}$ für jedes $L'' \| L$; hieraus folgt (j).

Jedes $\varkappa_{ij}$ erzeugt auf $\mathfrak{L}(i, j)$ eine zyklische Permutationsgruppe A_{ij}. Als abstrakte Gruppe ist sie bei endlicher Charakteristik von $\mathfrak{k}$ (s. II, §4, Ziffer 3) sehr leicht zu übersehen:

(k) *Es habe* $\mathfrak{k}$ *die Primzahlcharakteristik* p; *dann ist* p *die Ordnung der von* $\varkappa_{ij}$ $(i \neq j)$ *erzeugten Gruppe* A_{ij}.

Beweis. Da jede nichtidentische Translation von T die Ordnung p hat (vgl. II, Satz 4.5), folgt aus (1.6), daß $L^{\varkappa_{ij}^p} = L$ für jedes $L \in \mathfrak{L}(i, j)$ ist. Hieraus

folgt $x_{ij}^{p} = \underline{1}$ auf $\mathfrak{L}(i, j)$. Andererseits ist p auch die genaue Ordnung von x_{ij}, denn sonst wäre $x_{ij} = \underline{1}$, also $L^{x_{ij}} = L$ für jedes $L \in \mathfrak{L}(i, j)$, was wegen (h) für $i \neq j$ nicht der Fall sein kann.

(l) *Die Gruppe* A_{ijk} *der von* A_{ij} *und* A_{jk} *auf* $\mathfrak{L}(i, j, k) = \mathfrak{L}(i, j) \cap \mathfrak{L}(j, k)$ *erzeugten Permutationen ist abelsch.*

Hierzu brauchen wir nur die Gültigkeit von $x_{ij} x_{jk} = x_{jk} x_{ij}$ auf $\mathfrak{L}(i, j, k)$ zu beweisen: Sei $L \in \mathfrak{L}(i, j, k)$. Wir setzen $L^{x_{ij}} = L'$ und $L^{x_{jk}} = L^{\tau_{jk, L}} = L''$. Mit L gehört nach (a) auch L' zu $\mathfrak{L}(i, j, k)$. Mit (i) und (j) folgt, da T abelsch ist: $L^{x_{ij} x_{jk}} = L'^{x_{jk}} = L'^{\tau_{jk, L'}} = L'^{\tau_{jk, L}} = L^{\tau_{ij, L} \tau_{jk, L}} = L^{\tau_{jk, L} \tau_{ij, L}} = L''^{\tau_{ij, L}} = L''^{\tau_{ij, L''}} = L''^{x_{ij}} = L^{x_{jk} x_{ij}}$, was behauptet war.

Aus (l) folgt

$$\mathsf{A}_{ijk} = \mathsf{A}_{ij} \mathsf{A}_{ik} = \mathsf{A}_{ij} \mathsf{A}_{jk} = \mathsf{A}_{ik} \mathsf{A}_{jk}. \tag{1.7}$$

Ist insbesondere die Charakteristik von t eine Primzahl p, so ist wegen (k) offensichtlich A_{ijk} direktes Produkt zweier (zyklischer) Gruppen der Ordnung p.

(m) *Für eine endliche z. T.* t *gilt für* $i \neq j$ *stets* $\mathfrak{L}(i, j) \subset \mathfrak{L}$.

BEWEIS. Sei $\mathfrak{P} = \{L^{(0)}, \ldots, L^{(n)}\}$ das Büschel aller durch einen festen Punkt von $t = (T, \mathfrak{L})$ gehenden Geraden und $\mathfrak{P}(i, j) = \mathfrak{P} \cap \mathfrak{L}(i, j) = \{L^{(0)}, \ldots, L^{(m)}\}$. Da durch jeden Punkt von t mindestens drei Geraden gehen (s. III, Zusatz von Satz 1.4), folgt $n \geqq 2$. Aus $|L^{(r)}| = a_r$ $(r = 0, 1, \ldots, n)$ folgt $|T| = \sum_{r=0}^{n} a_r - n + 1$, nach (V) wegen [7]) $\mathfrak{P}^{x_{ij}} = \mathfrak{P}(i, j)^{x_{ij}}$ andererseits $|T| = \sum_{r=0}^{m} a_r$, also

$$\sum_{r=m+1}^{n} a_r = n - 1 \geqq 1. \tag{1.8}$$

Hieraus folgt notwendig $m < n$, d.h. $\mathfrak{P}(i, j) \subset \mathfrak{P}$, also auch $\mathfrak{L}(i, j) \subset \mathfrak{L}$, wie behauptet.

Aus (1.8) folgt wegen $a_r \geqq 2$ andererseits $2(n - m) \leqq n - 1$, wegen $m = |\mathfrak{P}(i, j)|$ und $n = |\mathfrak{P}|$ also

$$|\mathfrak{P}(i, j)| \geqq \tfrac{1}{2}(|\mathfrak{P}| + 1). \tag{1.9}$$

Das Gleichheitszeichen gilt dann und nur dann, wenn wirklich alle $a_r = 2$ sind; hierzu ist insbesondere Charakteristik 2 notwendig. Für Charakteristik $\neq 2$ können wir $a_r \geqq 3$ annehmen: dann folgt ebenso:

$$|\mathfrak{P}(i, j)| \geqq \tfrac{1}{3}(2|\mathfrak{P}| + 1). \tag{1.10}$$

(n) Die Bedingung (V) braucht nur für *ein* Geradenbüschel von $t = (T, \mathfrak{L})$ gefordert werden. Seien nämlich $p, p' \in T$ und $\mathfrak{P}$ bzw. $\mathfrak{P}'$ die Büschel mit p bzw. p' als Zentrum. Dann gibt es ein $\tau \in \mathsf{T}$ mit $p' = p^{\tau}$, also $\mathfrak{P}' = \mathfrak{P}^{\tau}$. Wegen (IV) ist nun $\mathfrak{P}'^{x_{ij}} = \mathfrak{P}^{\tau x_{ij}} = \mathfrak{P}^{x_{ij} \tau}$ für $i \neq j$ ebenfalls eine Zerlegung von t.

(o) Die Bedingung (VI) braucht nur für die Streckungen σ mit einem *festen* Zentrum gefordert zu werden. Seien nämlich $p, q \in T$, so gibt es ein

[7]) Siehe Fußnote 4.

$\tau \in \mathsf{T}$ mit $q = p^{\tau}$. Hat σ den Fixpunkt p, so läßt die Streckung $\tau^{-1}\sigma\tau$ den Punkt q fest, und mit (IV) erhalten wir $\tau^{-1}\sigma^{-1}\tau\varkappa_{ij}\tau^{-1}\sigma\tau = \tau^{-1}\sigma^{-1}\varkappa_{ij}\sigma\tau \neq \tau^{-1}\varkappa_{i\varrho_j\varrho}\tau = \varkappa_{i\varrho_j\varrho}$, wenn $\sigma^{-1}\varkappa_{ij}\sigma \neq \varkappa_{i\varrho_j\varrho}$ vorausgesetzt war.

(p) *Für jede affin-axiale Kollineation (d.h. Translation oder Streckung) π von $\mathfrak{k}$ genügen mit $\varkappa_{ij}$ auch alle $\pi^{-1}\varkappa_{ij}\pi$ den Bedingungen* (I) *bis* (VI) *einer kanonischen Abbildung.*

Beweis. Die Bedingungen (I) bis (III) und (VI) folgen sofort. (IV) schließt man so: Da mit τ auch $\pi\tau\pi^{-1}$ eine Translation ist (vgl. I, §4, Ziffer 4), ergibt sich, da (IV) für $\varkappa_{ij}$ gilt: $\pi\tau\pi^{-1}\varkappa_{ij} = \varkappa_{ij}\pi\tau\pi^{-1}$ und hieraus $\tau\pi^{-1}\varkappa_{ij}\pi = \pi^{-1}\varkappa_{ij}\pi\tau$, was behauptet war. Endlich folgt (V) daraus, daß mit $\mathfrak{P}^{\varkappa_{ij}}$ auch $\mathfrak{P}^{\varkappa_{ij}\pi}$ eine Zerlegung von $\mathfrak{k}$ ist.

§2. Die Konstruktion von T-Strukturen mit gegebenen kanonischen Abbildungen

1. In §1 sind wir, ausgehend von einer T-Struktur $\mathfrak{p}$, zu einer „Matrix" $(\varkappa_{ij})_{i,j\in I}$ von kanonischen Abbildungen $\varkappa_{ij}$ im Stamm $\mathfrak{k}$ von $\mathfrak{p}$ gelangt, welche den Eigenschaften (I) bis (VI) von Satz 1.2 genügen. Wir untersuchen jetzt das umgekehrte Problem: Gegeben sei eine z.T. $\mathfrak{k} = (T, \mathfrak{L})$ und eine Matrix $(\varkappa_{ij})$ von Abbildungen von $\mathfrak{L}$ auf sich oder der leeren Menge mit den Eigenschaften (I) bis (VI) (und damit auch deren Folgerungen); gesucht ist eine T-Struktur $\mathfrak{p}$, deren Stamm $\mathfrak{k}$ ist und deren kanonische Abbildungen (vgl. §1) mit den gegebenen $\varkappa_{ij}$ übereinstimmen. Eine solche T-Struktur existiert in der Tat immer; wir können somit jede T-Struktur durch ihren Stamm und ihre kanonischen Abbildungen eindeutig kennzeichnen.

2. Sei also $\mathfrak{k} = (T, \mathfrak{L})$ eine z.T., I eine Indexmenge und $(\varkappa_{ij})_{i,j\in I}$ eine Matrix von auf $\mathfrak{L}$ definierten Abbildungen, welche die Eigenschaften (I) bis (VI) haben. Wir denken uns $\mathfrak{k}$ in $|I|$ Exemplaren $\mathfrak{k}_i = (T_i, \mathfrak{L}_i)$ $(i \in I)$ gegeben, die paarweise zueinander disjunkt liegen sollen. Dabei entspreche dem Punkt p von $\mathfrak{k}$ auf $\mathfrak{k}_i$ der Punkt p_i. Die Abbildung $p \to p_i$ definiert uns einen Isomorphismus ι_i von $\mathfrak{k}$ auf $\mathfrak{k}_i$. Mittels $\mathfrak{k}$ und den $\varkappa_{ij}$ konstruieren wir jetzt eine Struktur $\mathfrak{p} = (P, \mathfrak{G}) = \{\mathfrak{k}, \varkappa_{ij}\}$, von der wir zeigen werden, daß sie eine über $\mathfrak{k}$ translationstreue T-Struktur mit genau $|I|$ Transitivitätsgebieten hinsichtlich ihrer Translationsgruppe T ist:

Als *Punkte* von $\mathfrak{p} = (P, \mathfrak{G})$ nehmen wir die Punkte aller $\mathfrak{k}_i$, d.h.

$$P = \bigcup_{i\in I} T_i; \tag{2.1}$$

als *Geraden* von $\mathfrak{p}$ sind genau die Punktmengen

$$L^i = \bigcup_{k\in I} L^{\varkappa_{ik}} \tag{2.2}$$

mit $L \in \mathfrak{L}$ und $i \in I$ definiert; zwei Geraden L^i und L'^j $(L, L' \in \mathfrak{L};\ i, j \in I)$ von $\mathfrak{p}$ heißen *parallel*, wenn L und L' in $\mathfrak{k}$ zueinander parallel sind.

Zur Definition (2.2) der Geraden von $\mathfrak{p}$ ist zu sagen, daß sie im folgenden Sinne unabhängig von der Wahl von $i \in I$ ist: *Zu jedem $j \in I$ mit $L^i \cap T_j \neq \emptyset$*

gibt es ein $L' \in \mathfrak{L}$ *mit* $L^i = L'^j$. Wegen $L^{\varkappa_{ij}} = \bigcup_{k \in I} L^{\varkappa_{ik}} \cap T_j = L^i \cap T_j \neq \emptyset$ folgt nämlich $L^{\varkappa_{ij}} \neq \emptyset$, mit (III) also

$$L^i = \bigcup_{k \in I} L^{\varkappa_{ik}\varkappa_{ik}} = \bigcup_{k \in I} L^{\varkappa_{ij}\varkappa_{jk}\varkappa_{ik}} = L'^j,$$

wenn $L' = L^{\varkappa_{ij}}$ gesetzt wird.

Auch die Parallelitätsdefinition ist unabhängig von der Wahl von i und j: *Aus* $L^i = L'^j$, $L^k_* = L'^k_*$ *und* $L^i \| L^k_*$ (d. h. $L \| L_*$) *folgt* $L' \| L'_*$. Soeben wurde nämlich $L' = L^{\varkappa_{ij}}$ bewiesen, ebenso gilt $L'_* = L^{\varkappa_{kk}}_*$. Nun folgt $L' \| L'_*$ sofort aus (I).

Somit sind die gegebenen Definitionen der Geraden und der Parallelität unabhängig von der besonderen Konstruktion, also durch t und die $\varkappa_{ij}$ eindeutig bestimmt.

3. Wir müssen nun zeigen, daß unsere so definierte Struktur $\mathfrak{p}$ tatsächlich alle gewünschten Eigenschaften hat. Zunächst gilt:

HILFSSATZ 2.1. *Die Struktur* $\{t, \varkappa_{ij}\}$ *ist eine a. P.*

BEWEIS. Wir müssen die eine a. P. kennzeichnenden Axiome nachprüfen: (a 0) und (a 1) haben wir bewiesen, wenn wir zeigen, daß durch zwei verschiedene Punkte p_i und q_j (aus T_i bzw. T_j) genau eine Gerade geht. Wir machen dazu die folgende Fallunterscheidung: (α) $i = j$: In diesem Fall ist $\langle p, q \rangle^i_t$ die gewünschte Gerade durch $p_i = p^{\varkappa_{ii}}$ und q_i, und eine weitere gibt es nicht, da $\langle p, q \rangle_t$ in t eindeutig bestimmt ist. (β) $i \neq j$: Sei $\mathfrak{P}$ das Büschel der durch p gehenden Geraden aus $\mathfrak{L}$. Nach (V) gibt es genau eine durch q gehende Gerade $L^{\varkappa_{ij}} \in \mathfrak{P}^{\varkappa_{ij}} (L \in \mathfrak{P})$. Die (eindeutig bestimmte) gewünschte Gerade ist L^i.

Axiom (a 2) beweist man so: Sei $L^i \in \mathfrak{G}$ und $p_j \in T_j$ $(i, j \in I)$. Aus der Definition der Parallelität folgt, daß $\langle L \| p \rangle^j_t$ eine zu L^i parallele durch p_j gehende Gerade aus $\mathfrak{G}$ ist; sie ist die einzige dieser Art, weil (a 2) in der z. T. t gilt. Die Parallelitätsrelation erfüllt ferner die drei Äquivalenzbedingungen, weil dasselbe für die Parallelität in t der Fall ist.

Da (a 3) in $t = (T, \mathfrak{L})$ gilt, gibt es drei Punkte $p, q, r \in T$, die nicht auf einer Geraden aus $\mathfrak{L}$ liegen. Wegen $L^i \cap T_i = L_i$ $(L \in \mathfrak{L})$ können dann auch $p_i, q_i, r_i \in T_i$ nicht auf einer Geraden aus $\mathfrak{G}$ liegen.

Schließlich folgt (a 4), sogar (t 2), nach der Geradendefinition (2.2) sofort aus der Gültigkeit von (t 2) für die Translationsstruktur t.

Da t zu allen t_i isomorph ist, können wir t als Unterstruktur von $\mathfrak{p} = \{t, \varkappa_{ij}\}$ auffassen. Wir wollen zeigen, daß $\mathfrak{p}$ translationstreu über t ist. Hierzu benötigen wir einige Hilfssätze:

HILFSSATZ 2.2. *Jede Translation* τ *von* t *induziert eine Translation* τ^* *auf* $\mathfrak{p}$, *die jedes* T_i *auf sich abbildet.*

Die durch $t^{\tau^*}_i = t^{\varkappa_{ii}\tau^*} = t^{\tau\varkappa_{ii}}$ $(t \in T)$ definierte Transformation bildet nämlich jedes T_i eineindeutig auf sich ab. Außerdem ist τ^* auf $\mathfrak{p}$ eine Translation; denn zunächst folgt aus $\tau \neq \underline{1}$ auf t, daß τ^* auf P keinen Punkt festläßt;

ist andererseits $\tau = \underline{1}$, so ist auch $\tau^* = \underline{1}$. Wegen (IV) gilt ferner

$$L^{i\tau^*} = \bigcup_{j \in I} L^{\varkappa_{ij}\, j\tau^*} = \bigcup_{j \in I} L^{\varkappa_{ij}\, \tau j} = \bigcup_{j \in I} L^{\tau \varkappa_{ij}\, j} = L^{\tau i} \| L^i,$$

letzteres nach unserer für $\mathfrak{p}$ gegebenen Definition der Parallelität.

Als Umkehrung von Hilfssatz 2.2 gilt:

HILFSSATZ 2.3. *Jede Translation τ von $\mathfrak{p}$ mit $T_i^\tau = T_i$ für ein bestimmtes $i \in I$ induziert eine Translation τ' auf $\mathfrak{t}$. Ferner gilt dann $T_j^\tau = T_j$ für alle $j \in I$.*

Wir setzen $\tau' = \iota_i \tau \iota_i^{-1}$, die so bestimmte Abbildung ist offensichtlich eine Translation auf $\mathfrak{t}$. Ferner erkennt man sofort, daß $\tau'^* = \tau$ und ebenso $\tau^{*\prime} = \tau$ ist (vgl. den Beweis von Hilfssatz 2.2). Hieraus folgt zusammen mit Hilfssatz 2.2 der zweite Teil unserer Behauptung.

Aus den beiden letzten Hilfssätzen folgt

HILFSSATZ 2.4. *Die Gruppe derjenigen Translationen von $\mathfrak{p}$, die ein (und damit jedes) T_i (als Ganzes) festlassen, ist zur Gruppe T aller Translationen von $\mathfrak{t}$ isomorph. Genau die T_i sind Transitivitätsgebiete hinsichtlich dieser Gruppe.*

BEWEIS. Die Abbildung $\tau \leftrightarrow \tau^*$ mit $\tau \in \mathsf{T}$ (vgl. den Beweis von Hilfssatz 2.2) vermittelt den gewünschten Isomorphismus. Da T auf T transitiv ist, so auch die Gruppe aller τ^* auf jedem T_i. Andere Transitivitätsgebiete dieser Gruppe gibt es nicht, da ja die T_i bereits ganz P überdecken.

BEMERKUNG. Wir können und wollen somit die beiden in Hilfssatz 2.4 vorkommenden Gruppen miteinander identifizieren und künftig T nennen. Wir setzen also $\tau^* = \tau' = \tau$.

HILFSSATZ 2.5. *Jedes $\tau \in \mathsf{T}$ ist zentral.*

BEWEIS. Bei vorgegebenem $\tau \in T$ sei $L \in \mathfrak{L}$ so gewählt, daß $L^\tau = L$ ist (ein solches L existiert, da $\mathfrak{t}$ zentral ist), dann gilt auch $L^{i\tau} = L^i$. Ferner sei $L'^i \| L^i$, nach unserer Parallelendefinition in $\mathfrak{p}$ (vgl. Ziffer 2) also auch $L' \| L$. Mit (IV) [vgl. auch den Beweis von Hilfssatz 2.2] folgt $L'^{i\tau} = L'^{\tau i} = L'^i$, letzteres, weil τ in $\mathfrak{t}$ zentral ist. Dies bedeutet aber, daß τ auch in $\mathfrak{p}$ zentral ist.

4. Wir geben als nächstes eine Übersicht über *alle* zentralen Translationen von $\mathfrak{p}$. Aus (VI) wird sich als Umkehrung von Hilfssatz 2.5 ergeben, daß jede zentrale Translation von $\mathfrak{p}$ zu T gehört. Im folgenden sei ζ stets eine zentrale Translation von $\mathfrak{p}$.

HILFSSATZ 2.6. *Jedes T_i^ζ $(i \in I)$ ist Transitivitätsgebiet hinsichtlich T.*

BEWEIS. Seien $p_i, q_i \in T_i$, dann gibt es ein $\tau \in \mathsf{T}$ mit $p_i^\tau = q_i$, denn T_i ist nach Hilfssatz 2.4 ein Transitivitätsgebiet bezüglich T. Nach I, Satz 4.5 ist τ mit ζ vertauschbar, wenn τ und ζ verschiedene Zentren haben. Stimmen die Zentren von τ und ζ überein, so wählen wir ein nichtidentisches $\tau' \in \mathsf{T}$, das ein anderes Zentrum als τ (und ζ) hat. Dann sind auch die Zentren von $\tau'\tau$ und ζ voneinander verschieden, und wir erhalten $\tau'\tau\zeta = \zeta\tau'\tau = \tau'\zeta\tau$. Jedenfalls gilt also stets $\tau\zeta = \zeta\tau$. Hieraus schließen wir $q_i^\zeta = p_i^{\tau\zeta} = p_i^{\zeta\tau}$, d.h. auch

p_i^ζ und q_i^ζ gehören ein und demselben Transitivitätsgebiet von T an; T_i^ζ ist also Transitivitätsgebiet hinsichtlich T.

Aus den Hilfssätzen 2.4 und 2.6 folgt, daß ζ eine Permutation $\bar{\zeta}$ auf der Menge aller T_i und damit auch auf I induziert: Wir können

$$T_i^\zeta = T_{i\bar{\zeta}} \tag{2.3}$$

setzen. Offensichtlich ist $\overline{\zeta\zeta'} = \bar{\zeta}\bar{\zeta}'$.

HILFSSATZ 2.7. *Jedes zentrale ζ permutiert die T_i regulär, d.h. die von ζ auf I induzierte Permutation $\bar{\zeta}$ ist regulär.*

BEWEIS. Aus der Existenz eines $i \in I$ und einer natürlichen Zahl n mit $i^{\bar{\zeta}^n} = i^{\bar{\zeta}^n} = i$ folgt $T_i^{\zeta^n} = T_i$, nach Hilfssatz 2.3 hieraus $T_j^{\zeta^n} = T_j$ für *alle* $j \in I$, d.h. $\bar{\zeta}^n = \underline{1}$, was behauptet war.

5. Es sei jetzt η eine (nicht notwendig zentrale) Translation von $\mathfrak{p}$, welche die T_i regulär permutiert, d.h. für die es eine auf I reguläre Permutation $\bar{\eta}$ gibt mit $T_i^\eta = T_{i\bar{\eta}}$. (Nach Hilfssatz 2.7 haben alle zentralen Translationen diese Eigenschaft.) Zur Abkürzung setzen wir $i^{\bar{\eta}} = i^{\bar{\eta}}$.

HILFSSATZ 2.8. *Die Abbildung $\pi_i = \iota_i \eta \iota_i^{-1}$ ist eine affin-axiale Kollineation (d.h. Translation oder Streckung) von $\mathfrak{t}$.*

BEWEIS. Zunächst ist klar, daß π_i eine eineindeutige Abbildung von T auf sich ist. Da η Translation ist, gilt ferner $L^i \| L^{i\eta} = L'^{\bar{\eta}}$; hierbei ist L' zu L parallel (vgl. die in Ziffer 2 gegebene Definition der Parallelität für $\mathfrak{p} = \{\mathfrak{t}, \varkappa_{ij}\}$). Also ist

$$L^{\pi_i} = L^{\iota_i \eta \iota_i^{-1}} = (L^i \cap T_i)^{\eta \bar{\iota}_i^{-1}} = (L^{i\eta} \cap T_i)^{\bar{\iota}_i^{-1}} = (L'^{\bar{\eta}} \cap T_i)^{\bar{\iota}_i^{-1}} = L_i'^{\bar{\iota}_i^{-1}} = L' \| L, \text{ w.z.b.w.}$$

HILFSSATZ 2.9. *Sei $\varrho \neq \underline{1}$ eine auf I reguläre Permutation und $j^\varrho = j^{\bar{\tau}}$ $(j \in I)$ gesetzt. Zu jedem j sei eine affin-axiale Kollineation π_j auf $\mathfrak{t}$ definiert. Die auf P definierte Abbildung*

$$\eta = \iota_j^{-1} \pi_j \iota_\tau, \tag{2.4}$$

ist genau dann eine Translation auf $\mathfrak{p}$, wenn alle π_j übereinstimmen (sie werden dann π genannt) und wenn

$$\varkappa_{\tau i, \tau j} = \pi^{-1} \varkappa_{ij} \pi \tag{2.5}$$

für alle $i, j \in I$ gilt.

BEWEIS. Jedenfalls ist η eineindeutig und wegen $\varrho \neq \underline{1}$ auf $\mathfrak{p}$ fixpunktfrei. Aus (2.4) folgt

$$L^{i\eta} = \bigcup_{j \in I} L^{\varkappa_{ij} \iota_j \eta} = \bigcup_{j \in I} L^{\varkappa_{ij} \pi_j \iota_j} = \bigcup_{j \in I} L^{\pi_j \pi_j^{-1} \varkappa_{ij} \pi_j \iota_j}$$

für alle $L \in \mathfrak{L}$. Damit dies eine Gerade (und damit dann eine zu L^i parallele) von $\mathfrak{p}$ ist, muß notwendig $L^{\pi_j \pi_j^{-1} \varkappa_{ij} \pi_j} = L'^{\varkappa_{\tau i, \tau j}}$ für ein $L' \in \mathfrak{L}$ und alle $i, j \in I$ gelten. Wegen (II) ergibt die Spezialisierung $j = i$ insbesondere $L' = L^{\pi_j}$. Für jedes $L \in \mathfrak{L}$ ist also L^{π_j} von j unabhängig, mithin hängt auch π_j von j nicht ab:

$\pi_j = \pi$. Nun folgt $\pi^{-1}\varkappa_{ij}\pi = \varkappa_{\bar{i}\bar{j}}$ als notwendige Bedingung dafür, daß η Translation ist. Sie ist aber auch hinreichend wegen

$$L^{i\eta} = \bigcup_{j \in I} L^{\pi \varkappa_{\bar{i}\bar{j}} \iota_{\bar{j}}^{-1}} = L^{\pi \bar{i}} \| L^i.$$

Es gilt die folgende Umkehrung:

HILFSSATZ 2.10. *Jede Translation η, welche die T_i regulär permutiert, läßt sich als Abbildung der in Hilfssatz 2.9 angegebenen Art darstellen, d.h. zu jedem solchen η gibt es eine affin-axiale Kollineation π von t mit $\eta = \iota_{\bar{j}}^{-1} \pi \iota_{\bar{j}}$ $(\bar{j} = j^{\bar{\eta}})$.*

BEWEIS. Nach Hilfssatz 2.8 ist jedenfalls $\iota_j \eta \iota_{\bar{j}}^{-1} = \pi_j$ für jedes $j \in I$ eine affin-axiale Kollineation auf t. Aus Hilfssatz 2.9 folgt, daß die π_j alle übereinstimmen. Setzen wir sie $= \pi$, so haben wir den gewünschten Sachverhalt.

Aus den Hilfssätzen 2.9 und 2.10 sowie aus (VI) folgt jetzt sofort

HILFSSATZ 2.11. *Jede Translation η, welche die T_i regulär permutiert, läßt jedes T_i fest, gehört also zu* T.

Mit Hilfssatz 2.7 folgt insbesondere der

ZUSATZ. *Jede zentrale Translation gehört zu* T.

Weitergehend zeigen wir jetzt

HILFSSATZ 2.12. *Jede Translation von $\mathfrak{p}$ ist zentral (und gehört mithin zu* T).

BEWEIS. Sei α eine Translation von $\mathfrak{p}$. Mit τ ist auch $\alpha^{-1}\tau\alpha$ eine zentrale Translation (vgl. I, §4, Ziffer 4), nach dem Zusatz zu Hilfssatz 2.11 gilt mithin $\alpha^{-1}\mathsf{T}\alpha = \mathsf{T}$. Jedes T_i^α ist Transitivitätsgebiet hinsichtlich T; denn es ist einerseits $T_i^{\alpha\mathsf{T}} = T_i^{\mathsf{T}\alpha} = T_i^\alpha$, andererseits bleibt aber auch keine echte Teilmenge U_i von T_i^α bei T fest. Wäre nämlich $U_i^{\mathsf{T}} = U_i$, so $U_i^{\alpha^{-1}\mathsf{T}} = U_i^{\alpha^{-1}} \subset T_i$; das ist aber nicht möglich, weil T_i ein Transitivitätsgebiet von T ist. Somit permutiert α die T_i untereinander; sie werden durch α sogar regulär permutiert: Aus $T_i^{\alpha^r} = T_i$ für ein $i \in I$ und eine natürliche Zahl r folgt nämlich $T_j^{\alpha^r} = T_j$ für alle $j \in I$ nach Hilfssatz 2.3. Wegen Hilfssatz 2.11 gehört also α zu T, und α ist zentral.

HILFSSATZ 2.13. *Auf $\mathfrak{p} = (P, \mathfrak{G})$ gibt es Translationen in jeder Richtung, d.h. zu jedem $G \in \mathfrak{G}$ gibt es ein $\tau \neq \underline{1}$ aus* T *mit $G^\tau = G$.*

BEWEIS. Zu jedem $G \in \mathfrak{G}$ gibt es ein $L \in \mathfrak{L}$ und ein $i \in I$ mit $G = L^i$ [vgl. die Definition (2.2)]. Weil t eine z.T. ist, gibt es ein $\tau \neq \underline{1}$ aus T mit $L^\tau = L$. Für dasselbe τ gilt auch $G^\tau = G$.

6. Nunmehr haben wir alle Hilfsmittel zur Hand, um den Hauptsatz dieses Paragraphen beweisen zu können, welcher die Ergebnisse von §1 umkehrt.

SATZ 2.1. *Zu jeder zentralen Translationsstruktur t und jeder Matrix kanonischer Abbildungen $\varkappa_{ij}$ $(i, j \in I)$ über t [mit den Eigenschaften* (I) *bis* (VI) *von Satz* 1.2] *gibt es eine Struktur $\mathfrak{p} = \{t, \varkappa_{ij}\}$ (die nach dem in Ziffer* 2 *beschriebenen Verfahren konstruiert wird). Sie ist eine über t translationstreue T-Struktur.*

BEWEIS. Nach Hilfssatz 2.1 ist $\mathfrak{p}$ eine a.P. Nach den Hilfssätzen 2.2 und 2.12 ist $\mathfrak{p}$ zentral und translationstreu über t. Wegen Hilfssatz 2.13 ist $\mathfrak{p}$ sogar eine T-Struktur.

§ 3. Weitere Eigenschaften der T-Strukturen

1. In §1 sind wir von einer T-Struktur $\mathfrak{p}$ ausgegangen und haben dazu eine z.T., den Stamm $\mathfrak{l}=\mathfrak{l}_{\mathfrak{p}}$ von $\mathfrak{p}$, und eine Schar kanonischer Abbildungen $\varkappa_{ij}=\varkappa_{ij}(\mathfrak{p})$ mit gewissen Eigenschaften gefunden. In §2 gingen wir umgekehrt von einer z.T. $\mathfrak{l}$ und kanonischen Abbildungen $\varkappa_{ij}$ aus und gelangten zu einer T-Struktur $\mathfrak{p}=\{\mathfrak{l}, \varkappa_{ij}\}$. Wir müssen jetzt zeigen, daß beide Konstruktionen hintereinander ausgeführt wieder zum Ausgangsbereich zurückführen. Genauer gesagt müssen wir also die Gültigkeit von

$$\mathfrak{l}_{\{\mathfrak{l},\varkappa_{ij}\}}=\mathfrak{l}, \qquad \varkappa_{ij}(\{\mathfrak{l},\varkappa_{ij}\})=\varkappa_{ij} \tag{3.1}$$

und

$$\{\mathfrak{l}_{\mathfrak{p}}, \varkappa_{ij}(\mathfrak{p})\}=\mathfrak{p} \tag{3.2}$$

nachweisen (wobei isomorphe Bereiche in üblicher Weise identifiziert werden, wie im folgenden ausgeführt wird).

BEWEIS VON (3.1). Da $\{\mathfrak{l}, \varkappa_{ij}\}$ translationstreu über $\mathfrak{l}$ (vgl. Satz 2.1), also $\mathfrak{l}$ Stamm von $\{\mathfrak{l}, \varkappa_{ij}\}$ ist (vgl. III, §2), gilt $\mathfrak{l}_{\{\mathfrak{l},\varkappa_{ij}\}}=\mathfrak{l}$. Die zweite Gleichung von (3.1) folgt mit Hilfe von (2.2) und (1.1) so:

$$\bigcup_{j\in I} L^{\varkappa_{ij}\eta}=L^i=\bigcup_{j\in I}(L^i\cap T_j)=\bigcup_{j\in I}(L^i\cap T_j)^{\varphi^{-1}\eta}$$

$$=\bigcup_{j\in I} L^{\varkappa_{ij}((\mathfrak{l},\varkappa_{ij}))\eta}, \quad \text{also} \quad L^{\varkappa_{ij}}=L^{\varkappa_{ij}((\mathfrak{l},\varkappa_{ij}))}$$

für jedes $L\in\mathfrak{L}$ und $i, j\in I$.

BEWEIS VON (3.2). Seien T bzw. T* die Translationsgruppen von $\mathfrak{p}=(P, \mathfrak{G})$ bzw. $\{\mathfrak{l}_{\mathfrak{p}}, \varkappa_{ij}(\mathfrak{p})\}=\mathfrak{p}^*=(P^*, \mathfrak{G}^*)$. Die Mächtigkeiten der Mengen der Transitivitätsgebiete von P bzw. P^* hinsichtlich T bzw. T* sind $|I|$, stimmen also überein. Ebenso haben $\mathfrak{p}$ und $\mathfrak{p}^*$ (bis auf Isomorphie) den gleichen Stamm $\mathfrak{l}_{\mathfrak{p}}$. Wir können und wollen somit die Transitivitätsgebiete T_i und T_i^* für jedes $i\in I$, mithin auch P und P^*, miteinander identifizieren, ebenfalls die Translationsgruppen T und T*, die ja isomorph zur Translationsgruppe von $\mathfrak{l}_{\mathfrak{p}}$ sind.

Sei L eine Gerade von $\mathfrak{l}_{\mathfrak{p}}$ und L^i bzw. L^{i*} diejenige Gerade aus $\mathfrak{G}$ bzw. $\mathfrak{G}^*$, welche in T_i die Spur $L_i=L^\eta$ hat. Mit Hilfe von (2.2) und (1.1) erhalten wir nun

$$L^{i*}=\bigcup_{j\in I} L^{\varkappa_{ij}(\mathfrak{p})\eta}=\bigcup_{j\in I}(L^i\cap T_j)^{\varphi^{-1}\eta}=\bigcup_{j\in I}(L^i\cap T_j)=L^i.$$

Auch die Geraden und die Inzidenzbeziehungen stimmen also in $\mathfrak{p}$ und $\mathfrak{p}^*$ überein.

Sind L^i und L'^j ($i, j\in I$; L und L' Geraden von $\mathfrak{l}_{\mathfrak{p}}$) in $\mathfrak{p}$ parallel, so ist $L\|L'$ in $\mathfrak{l}_{\mathfrak{p}}$, hieraus folgt aber die Parallelität von L^i und L'^j auch in $\mathfrak{p}^*$, ebenso gilt die Umkehrung. Damit ist gezeigt, daß $\mathfrak{p}$ mit $\mathfrak{p}^*$ übereinstimmt.

2. Wir können nunmehr die bisherigen Ergebnisse von Teil IV in dem folgenden Hauptsatz zusammenfassen, der eine gewisse Übersicht über den Bau der T-Strukturen gibt.

SATZ 3.1. *Zu jeder zentralen Translationsstruktur t und jeder Matrix $(\varkappa_{ij})_{i,j\in I}$ von kanonischen Abbildungen [die den Bedingungen* (I) *bis* (VI) *von Satz* 1.2 *genügen] kann man eine T-Struktur $\mathfrak{p}=\{t, \varkappa_{ij}\}$ mit genau $|I|$ Transitivitätsgebieten hinsichtlich ihrer Translationsgruppe bilden: Die Punkte und Geraden von $\mathfrak{p}$ werden durch* (2.1) *bzw.* (2.2) *definiert, die Parallelität zwischen Geraden wird als naturgemäße Erweiterung der in t gegebenen Parallelität eingeführt [s. die Definition im Anschluß an* (2.2)]. *Die T-Struktur $\mathfrak{p}$ ist über t translationstreu. Umgekehrt läßt sich jede über t translationstreue T-Struktur in der oben angegebenen Gestalt darstellen. Die so erhaltene Zuordnung zwischen den über t translationstreuen T-Strukturen einerseits und den Strukturen $\{t, \varkappa_{ij}\}$ andererseits ist umkehrbar eindeutig.*

3. Der soeben angegebene Struktursatz gestattet uns nun, nichttriviale Beispiele von T-Strukturen anzugeben und weitere Eigenschaften herzuleiten. Zunächst wollen wir uns aber eine Übersicht über die Kollineationen einer T-Struktur $\mathfrak{p}=\{t, \varkappa_{ij}\}$ verschaffen. Wir formulieren zuerst den folgenden, der Folgerung (p) von §1 (Ziffer 5) analogen Satz, der auch entsprechend wie (p) bewiesen wird.

SATZ 3.2. *Für jede Kollineation γ von t genügen mit $\varkappa_{ij}$ auch alle $\gamma^{-1}\varkappa_{ij}\gamma$ den Bedingungen* (I) *bis* (VI), *sind also kanonische Abbildungen.*

Sei nun γ eine Kollineation von $\mathfrak{p}$. Zunächst beweisen wir zwei Hilfssätze über γ.

HILFSSATZ 3.1. *Jedes T_i^γ $(i\in I)$ ist Transitivitätsgebiet hinsichtlich* T. *Mithin induziert γ auf I eine Permutation $\bar\gamma$ durch*

$$T_i^\gamma = T_{i^{\bar\gamma}}.$$

BEWEIS. Seien $p_i, q_i\in T_i$; es gibt ein $\tau\in\mathsf{T}$ mit $q_i=p_i^\tau$, also gilt $q_i^\gamma=p_i^{\tau\gamma}=p_i^{\gamma\gamma^{-1}\tau\gamma}$. Wegen $\gamma^{-1}\tau\gamma\in\mathsf{T}$ ist alles gezeigt.

Analog wie in §2 setzen wir $i^{\bar\gamma}=\bar\imath$ zur Abkürzung.

HILFSSATZ 3.2. *Die Abbildung $\gamma_i=\iota_i\gamma\iota_{\bar\imath}^{-1}$ ist eine Kollineation in t.*

Der Beweis geht entsprechend wie der von Hilfssatz 2.8: Zunächst ist γ_i eineindeutig auf T. Sei $L\in\mathfrak{L}$, dann ist $L^{\gamma_i}=L^{\iota_i\gamma\iota_{\bar\imath}^{-1}}=(L^{\iota_i}\cap T_i)^{\gamma\iota_{\bar\imath}^{-1}}=(L^{\iota_i\gamma}\cap T_{\bar\imath})^{\iota_{\bar\imath}^{-1}}\in\mathfrak{L}$ wegen $L^{\iota_i\gamma}\in\mathfrak{G}$. Aus $L\parallel L'$ folgt $L^{\iota_i}\parallel L'^{\iota_j}$ $(i,j\in I)$, hieraus $L^{\iota_i\gamma}\parallel L'^{\iota_i\gamma}$ und daraus schließlich $L^{\gamma_i}\parallel L'^{\gamma_i}$. Damit sind alle Eigenschaften einer Kollineation gezeigt.

Nunmehr können wir den Satz über die Kollineationen einer T-Struktur formulieren und beweisen:

SATZ 3.3. *Jede Kollineation γ der T-Struktur $\mathfrak{p}=\{t, \varkappa_{ij}\}$ induziert auf t eine Kollineation γ^* und auf I eine Permutation $\bar\gamma$ derart, daß*

$$\varkappa_{\bar\imath\bar\jmath}=\gamma^{*-1}\varkappa_{ij}\gamma^* \qquad (i,j\in I;\ i^{\bar\gamma}=\bar\imath,\ j^{\bar\gamma}=\bar\jmath) \tag{3.3}$$

und

$$t_j^\gamma = t^{\gamma^*\iota_j} \qquad (t\in T) \tag{3.4}$$

wird. Umgekehrt definieren γ^* und $\bar\gamma$, wenn sie durch (3.3) *miteinander verknüpft sind, durch* (3.4) *eine Kollineation γ auf $\mathfrak{p}$.*

BEWEIS. Nach Hilfssatz 3.2 ist $\gamma_j = \iota_j \gamma \iota_j^{-1}$ auf t eine Kollineation, mithin $L^{\gamma_j} \in \mathfrak{L}$, wenn $L \in \mathfrak{L}$ ist. Wir erhalten

$$L^{i\gamma} = \bigcup_{j \in I} L^{x_{ij}\gamma_j\gamma} = \bigcup_{j \in I} L^{x_{ij}\gamma_j\gamma_j} = \bigcup_{j \in I} L^{\gamma_j\gamma_j^{-1}x_{ij}\gamma_j\gamma_j}.$$

Dieser Ausdruck soll eine Gerade darstellen, er ist mithin $= \bigcup_{j \in I} L'^{x_{ij}\gamma_j}$ für ein $L' \in \mathfrak{L}$. Hieraus folgt

$$L^{\gamma_j\gamma_j^{-1}x_{ij}\gamma_j} = L'^{x_{ij}\gamma_j} \tag{3.5}$$

für alle $i, j \in I$. Für $i = j$ erhalten wir wegen (II) insbesondere $L' = L^{\gamma_j}$. Da L' nicht von j abhängt, stimmen alle γ_j überein: $\gamma_j = \gamma^*$ für jedes $j \in I$; das ergibt $\iota_j \gamma \iota_j^{-1} = \gamma^*$, also (3.4). Außerdem folgt (3.3) jetzt sofort aus (3.5). Sind umgekehrt (3.3) und (3.4) erfüllt, so erhalten wir

$$L^{i\gamma} = \bigcup_{j \in I} L^{\gamma^* x_{ij}\gamma_j} = L^{\gamma^*\bar{\gamma}},$$

und man sieht nun leicht, daß γ eine Kollineation ist.

4. Wir wollen jetzt die Streckungen von $\{t, x_{ij}\}$ untersuchen. Sei σ eine Streckung mit dem Zentrum $t_i \in T_i$ $(i \in I)$, und σ^* bzw. $\bar{\sigma}$ seien entsprechend wie γ^*, $\bar{\gamma}$ in der vorigen Ziffer definiert; ferner sei $j^{\bar{\sigma}} = \bar{j}$ $(j \in I)$ gesetzt.

HILFSSATZ 3.3. *Die Kollineation $\sigma^* = \iota_i \sigma \iota_i^{-1}$ ist eine Streckung in t mit dem Fixpunkt t.*

BEWEIS. Nach Hilfssatz 3.2 ist σ^* jedenfalls eine Kollineation. Wegen $L^{i\sigma} \| L^i$ ist auch $L^{\sigma^*} \| L$. Da t_i Fixpunkt von σ ist, bleibt $t \in T$ bei σ^* fest.

HILFSSATZ 3.4. *Die Permutation $\bar{\sigma}$ läßt i fest und permutiert die von i verschiedenen Elemente aus I regulär.*

BEWEIS. Aus $t_i^{\sigma} = t_i$ und Hilfssatz 3.1 folgt $i^{\bar{\sigma}} = i$. Sei n eine natürliche Zahl. Ist $j^{\bar{\sigma}^n} = j^{\overline{\sigma^n}} = j$ für ein $j \neq i$, so hat σ^n die beiden voneinander verschiedenen Fixpunkte t_i und t_j; denn es ist $\sigma^{*n} = \iota_j \sigma^n \iota_j^{-1}$, also $t_j^{\sigma^n} = t^{\sigma^{*n}\iota_j} = t^{\iota_j} = t_j$. Das bedeutet aber (vgl. I, Satz 4.1) $\sigma^n = \underline{1}$, also auch $\bar{\sigma}^n = \underline{1}$.

Aus Satz 3.3 folgt für Streckungen die folgende Spezialisierung:

SATZ 3.4. *Jede Streckung σ der T-Struktur $\mathfrak{x} = \{t, x_{ij}\}$ mit Fixpunkt in T_i $(i \in I)$ induziert auf t eine Streckung σ^* und auf I eine i festlassende auf $I - \{i\}$ reguläre Permutation $\bar{\sigma}$ derart, daß*

$$x_{j^{\bar{\sigma}}k^{\bar{\sigma}}} = \sigma^{*-1} x_{jk} \sigma^* \qquad (j, k \in I;\ j^{\bar{\sigma}} = \bar{j},\ k^{\bar{\sigma}} = \bar{k}) \tag{3.6}$$

und

$$t_j^{\sigma} = t^{\sigma^*\iota_j} \qquad (t \in T) \tag{3.7}$$

ist. Umgekehrt definieren σ^* und $\bar{\sigma}$, wenn sie vermöge (3.6) *miteinander verknüpft sind, durch* (3.7) *eine Streckung σ mit Fixpunkt in T_i.*

BEWEIS. (a) Sei σ eine Streckung mit Fixpunkt in T_i; da σ eine Kollineation ist, können wir Satz 3.3 anwenden, insbesondere gemäß dieses Satzes σ^* und $\bar{\sigma}$ bilden. Nach Hilfssatz 3.3 ist σ^* eine Streckung auf t, nach Hilfssatz 3.4

permutiert $\bar{\sigma}$ die Elemente von $I-\{i\}$ regulär, während i festbleibt. Aus (3.3) bzw. (3.4) folgt (3.6) bzw. (3.7).

(b) Sei jetzt umgekehrt σ^* eine Streckung auf t und $\bar{\sigma}$ eine auf $I-\{i\}$ reguläre Permutation, die i festläßt; ferner sei (3.6) erfüllt. Dann ist das durch (3.7) definierte σ nach Satz 3.3 jedenfalls eine Kollineation. Um zu zeigen, daß σ eine Streckung ist, bemerken wir, daß wegen $L^{\sigma^*} \| L$ auch $L^{j\sigma} = L^{\sigma^* j}$ zu L^j parallel ist; außerdem hat σ einen Fixpunkt in T_i; ist nämlich $t \in T$ Fixpunkt von σ^*, so gilt $t_i^\sigma = t^{\sigma^* i} = t_i = t_i \in T_i$.

Satz 3.4 gestattet eine interessante Folgerung. Nach ihm ist nämlich auch die durch

$$t_j \to t_{j\varrho} \qquad (t \in T,\ j \in I) \tag{3.8}$$

definierte Abbildung eine Streckung von $\mathfrak{p}$, wenn ϱ eine Permutation auf I ist, welche i festläßt und die übrigen $j \in I$ regulär permutiert, und wenn ferner $\varkappa_{j\varrho\, k\varrho} = \varkappa_{jk}$ für alle $j, k \in I$ gilt. Da aber die Abbildung (3.8) alle Punkte aus T_i festläßt, ist die so definierte Streckung die Identität; es muß also auch notwendig $\varrho = \underline{1}$ sein.

Die Negation dieser Aussage ergibt einen Sachverhalt, der ähnlich wie die Bedingung (VI) aussieht:

Satz 3.5. *Zu jedem $i \in I$ und jeder Permutation $\varrho \neq \underline{1}$ auf I, welche i festläßt und die übrigen Elemente von I regulär permutiert, gibt es $j, k \in I$ mit*

$$\varkappa_{j\varrho\, k\varrho} \neq \varkappa_{jk}. \tag{3.9}$$

5. Wir nehmen nun an, daß $\mathfrak{p} = \{t, \varkappa_{ij}\}$ *endlich* sei. In diesem Falle kann $I = \{0, 1, \ldots, t\}$ gesetzt werden; die Anzahl der Transitivitätsgebiete hinsichtlich T ist dann $t+1$. Aus I, Satz 5.3 folgt, daß in höchstens einem Transitivitätsgebiet T_i $(i = 0, \ldots, t)$ Zentren nichtidentischer Streckungen liegen. O.B.d.A. sei dies T_0, d.h. aus $p \notin T_0$ folgt $\Pi_p = \{\underline{1}\}$. Wir nennen in diesem Falle T_0 den *Sockel* von $\mathfrak{p}$. Die übrigen T_i werden durch eine Streckung aus Π_{p_0} $(p_0 \in T_0)$ regulär permutiert (vgl. Satz 3.4). Sei q_0 ein weiterer Punkt aus T_0; dann gibt es ein (eindeutig bestimmtes) $\tau \in \mathsf{T}$ mit $q_0 = p_0^\tau$. Aus I, Satz 5.3 (a) folgt $\Pi_{q_0} = \tau^{-1} \Pi_{p_0} \tau$; die Streckungsgruppen mit Zentren in T_0 sind also alle zueinander isomorph, und zwar nicht nur als abstrakte Gruppen, sondern auch als Permutationsgruppen; wir können nun von *der* Streckungsgruppe Π_* von $\mathfrak{p}$ reden, die als zu einer (und damit zu allen) Π_{p_0} mit $p_0 \in T_0$ isomorph erklärt wird.

Satz 3.6. *Die Streckungsgruppe Π_* einer endlichen T-Struktur ist Untergruppe der multiplikativen Gruppe eines Galois-Feldes, insbesondere also zyklisch*[8]).

Beweis. Seien $\sigma, \sigma' \in \Pi_{p_0}$, ferner $\sigma^* = \iota_j \sigma \iota_j^{-1}$ und $\sigma'^* = \iota_j \sigma' \iota_j^{-1} = \iota_i \sigma \sigma' \iota_j^{-1} \sigma$ (wie in Hilfssatz 3.3 eingeführt). Es folgt $(\sigma\sigma')^* = \sigma^* \sigma'^*$. Ist $\sigma^* = \underline{1}$, so $\iota_j \sigma = \iota_i \sigma$, also $q_j^\sigma = q_i \sigma$ für jedes $q \in T$; insbesondere läßt also σ jeden Punkt von T_0 fest, es ist also $\sigma = \underline{1}$. Somit ist Π_{p_0}, also auch Π_* zur Gruppe Π_p^* aller σ^* isomorph.

[8]) Dieser Satz gilt bereits für endliche t-Strukturen und kann hierfür mit ähnlichen Methoden bewiesen werden.

Die letztere ist aber Untergruppe der Gruppe Π_p aller p-Streckungen von t. Nun ist aber Π_p multiplikative Gruppe eines Galois-Feldes, nämlich des Kernes von t (vgl. II, Sätze 4.2 und 4.4).

Sei ϱ eine auf $I-\{i\}$ reguläre Permutation und $i^\varrho = i$, ferner σ^* eine Streckung auf t. Gilt $\varkappa_{j^\varrho k^\varrho} = \sigma^{*-1} \varkappa_{jk} \sigma^*$ für alle $j, k \in I$, so definiert nach Satz 3.4 die Abbildung $p_j \to p^{\sigma^* u^\varrho}$ eine Streckung σ auf $\mathfrak{p}$ mit einem Fixpunkt in T_i. Nach dem vorhin bemerkten kann $\sigma \neq \underline{1}$ für höchstens ein $i \in I$ gelten (nämlich nur für den Sockel T_i von $\mathfrak{p}$). Verneinen wir diese Aussage, so erhalten wir

SATZ 3.7. *Sei $\mathfrak{p} = \{t, \varkappa_{ij}\}$ eine endliche T-Struktur, σ^* eine Streckung auf t und ϱ eine i festlassende und auf $I-\{i\}$ reguläre Permutation, ferner seien σ^* und ϱ nicht beide $= \underline{1}$. Zu jedem solchen σ^*, zu jedem $i \in I$ mit höchstens einer Ausnahme und jedem auf $I-\{i\}$ regulärem ϱ gibt es dann $j, k \in I$ mit*

$$\varkappa_{j^\varrho k^\varrho} \neq \sigma^{*-1} \varkappa_{jk} \sigma^*. \tag{3.10}$$

§ 4. Beispiele von T-Strukturen

1. Wir wollen jetzt Beispiele von T-Strukturen, d.h. von Strukturen der Art $\{t, \varkappa_{ij}\}$ aufstellen, die keine Translationsstrukturen sind. Da $\mathfrak{P}^{\varkappa_{ij}}$ für $i \neq j$ nach (V) eine Zerlegung von t ist, müssen wir uns zunächst vergewissern, ob solche Zerlegungen überhaupt existieren können. Das ist in der Tat der Fall. Wir werden hier keine Übersicht über alle Zerlegungen einer vorgegebenen z.T. t geben, sondern uns mit einigen Beispielen begnügen, die zeigen sollen, daß die Untersuchungen über T-Strukturen nicht zu leeren Aussagen führen.

Um zu Beispielen von Zerlegungen einer z.T. zu gelangen, führen wir zunächst den Begriff einer Zerlegung in einem K-Rechtsmodul T ein. (Hierbei sei K ein Schiefkörper.) Ein *linearer Unterraum* von T ist eine Nebenklasse nach einem K-Untermodul $U \neq \{o\}$[9]) von T, d.h. eine Menge der Form $U+t$ mit $t \in T$. Zwei Unterräume $U+t$ und $U'+t'$ heißen zueinander windschief, wenn sie einen leeren Durchschnitt und außerdem U und U' nur das Nullelement o von T gemeinsam haben[10]), d.h. wenn $(U+t) \cap (U'+t') = \emptyset$ und $U \cap U' = \{o\}$ ist. Die Abbildungen $t \to t+a$ $(a \in T)$ sind die *Translationen* von T; sie bilden eine zu T isomorphe Gruppe T. Als eine *Zerlegung* von T definieren wir nun eine Menge paarweise windschiefer Unterräume von T, die ganz T überdecken.

Wir zeigen zunächst, daß jede Zerlegung $\mathfrak{Z}$ eines K-Moduls T auch eine Zerlegung einer geeignet gewählten z.T. ist, deren Punkte die Elemente von T sind, und deren Translationsgruppe mit der von T übereinstimmt. Zu diesem Zwecke konstruieren wir zu unserem $\mathfrak{Z}$ eine Partition $\mathfrak{P}$ des Moduls T

[9]) Wir schließen also hier, im Gegensatz zur üblichen Definition der linearen Unterräume, Punkte ausdrücklich aus.

[10]) Das hat folgende Bewandnis: Man schließe in üblicher Weise T projektiv ab, indem man als Punkte zu den Elementen von T noch die eindimensionalen Unterräume t K $(t \in T)$ als „uneigentliche Punkte" hinzunimmt und t K genau dann zum Unterraum $U+t'$ gehören soll, wenn t K $\subseteq U$ gilt. Dann zeigt sich, daß zwei lineare Unterräume genau dann zueinander windschief liegen, wenn ihr Durchschnitt auch bei Hinzunahme ihrer uneigentlichen Punkte leer bleibt.

auf folgende Weise: Zu $\mathfrak{P}$ sollen zunächst alle diejenigen Untermoduln U von T gehören, zu denen es ein $t \in T$ mit $U + t \in \mathfrak{Z}$ gibt. Je zwei verschiedene Untermoduln dieser Art haben nur das Nullelement von T gemeinsam, weil die Räume von $\mathfrak{Z}$ paarweise windschief liegen. Ferner sollen alle diejenigen einrangigen Untermoduln $t\mathsf{K}$ ($t \in T$) von T zu $\mathfrak{P}$ gehören, die in keinem der eben konstruierten U liegen. Die so gewonnene Menge $\mathfrak{P}$ von Untermoduln ist offensichtlich eine Partition von T, und die Struktur $\mathfrak{t} = [T, \mathfrak{P}]$ (vgl. II, §1) unsere gesuchte z.T.[11]; denn $\mathfrak{Z}$ ist offensichtlich Zerlegung von $\mathfrak{t}$ (im Sinne unserer in §1, Ziffer 3, gegebenen Definition).

Um Beispiele von Zerlegungen von z.T. zu finden, genügt es also, Zerlegungen von K-Moduln T zu suchen. Der Einfachheit halber wollen wir hier annehmen, daß T über K einen endlichen Rang d habe. Die Menge aller linearen Unterräume von T bilden den d-dimensionalen affinen Raum $\mathfrak{A}$ über K. Durch Hinzufügung der uneigentlichen Hyperebene H erweitern wir $\mathfrak{A}$ zum d-dimensionalen projektiven Raum $\mathfrak{R}$. In $\mathfrak{R}$ bilden wir eine Partition $\mathfrak{Q}$, d.h. eine Menge paarweise fremd liegender Unterräume von $\mathfrak{R}$, die ganz $\mathfrak{R}$ überdecken. Für $d > 2$ ist dies im allgemeinen auf mannigfache Art möglich [z.B. durch Kongruenzen im Sinne von ANDRÉ (1954), §8][12]. Aus $\mathfrak{Q}$ können wir jetzt sofort eine Zerlegung $\mathfrak{Z} = \bar{\mathfrak{Q}}$ in $\mathfrak{A}$ bilden: Als Unterräume aus $\bar{\mathfrak{Q}}$ nehmen wir die Mengen $\bar{Q} = Q - (Q \cap H)$ mit $Q \in \mathfrak{Q}$; man sieht sofort, daß für $\bar{\mathfrak{Q}}$ alle Bedingungen für eine Zerlegung erfüllt sind.

2. Nunmehr wollen wir Beispiele von T-Strukturen, die keine Translationsstrukturen sind, wirklich angeben. Nach den Untersuchungen der §§1 bis 3 müssen wir hierfür zu einer z.T. $\mathfrak{t}$ kanonische Abbildungen $\varkappa_{ij}$ finden, die den Eigenschaften (I) bis (VI) von Satz 1.2 genügen. Die Hauptschwierigkeit liegt im allgemeinen im Beweis von (VI).

Beispiele unendlicher T-Strukturen (die sogar affine Ebenen sind) finden sich in ANDRÉ (1959). Hier wollen wir endliche T-Strukturen angeben, die allerdings keine affinen Ebenen sind. Da es eine T-Struktur mit genau zwei Transitivitätsgebieten hinsichtlich ihrer Translationsgruppe nicht geben kann (wie im nächsten Paragraphen bewiesen wird), muß dessen Anzahl mindestens drei sein. Wir werden jetzt endliche T-Strukturen der Charakteristik p mit genau drei Transitivitätsgebieten konstruieren. Um mit unserem Verfahren zum Ziel zu kommen, müssen wir $p \geqq 5$ voraussetzen.

Sei $\mathfrak{t} = (T, \mathfrak{L})$ eine z.T., in der eine Zerlegung $\mathfrak{Z}$ möglich ist. (Beispiele hierfür haben wir in der vorigen Ziffer angegeben.) Dies $\mathfrak{t}$ soll der Stamm unserer gesuchten T-Struktur $\mathfrak{p}$ werden. Sei $\mathfrak{P}$ das Büschel aller durch einen festen Punkt $t \in T$ gehenden Geraden aus $\mathfrak{L}$ und $\varkappa$ eine Abbildung, die auf den Geraden von $\mathfrak{P}$ in folgender Weise definiert ist: $L^{\varkappa}$ sei diejenige Gerade aus $\mathfrak{Z}$, die zu $L \in \mathfrak{P}$ parallel ist (es gibt höchstens eine solche Gerade); existiert keine Gerade dieser Art, so sei $L^{\varkappa} = \emptyset$ gesetzt. Wir erweitern jetzt $\varkappa$ auf ganz $\mathfrak{L}$

[11] Dies ist nur eine der Möglichkeiten, aus einer Zerlegung $\mathfrak{Z}$ eines K-Rechtsmoduls T eine z.T. mit derselben Punktmenge T und derselben Zerlegung zu finden.

[12] Dagegen gibt es für $d = 2$ keine solche Möglichkeit.

vermöge folgender Definition: Zu jedem $L \in \mathfrak{L}$ gibt es ein $L' \in \mathfrak{P}$ und eine Translation $\tau \in \mathsf{T}$ mit $L'^{\tau} = L$; hierbei ist $\tau \bmod \mathsf{T}(L')$ eindeutig bestimmt. Wir setzen

$$L^{\varkappa} = L'^{\tau\varkappa} = L'^{\varkappa\tau}; \tag{4.1}$$

hierdurch ist $L^{\varkappa}$ eindeutig bestimmt. Ferner setzen wir $\emptyset^{\varkappa} = \emptyset$. Die Abbildung $\varkappa$ ist eineindeutig auf der Menge $\mathfrak{L}_0$ aller $L \in \mathfrak{L}$ mit $L^{\varkappa} \neq \emptyset$ ($\mathfrak{L}_0$ ist offensichtlich die Menge derjenigen Geraden, die zu denen der Zerlegung parallel sind), mithin können wir $\varkappa^{-1}$ auf $\mathfrak{L}_0$ als die zu $\varkappa$ inverse Abbildung erklären; für $L \notin \mathfrak{L}_0$, d.h. $L^{\varkappa} = \emptyset$, setzen wir auch $L^{\varkappa^{-1}} = \emptyset$, ferner erweitern wir $\varkappa^{-1}$ durch die Festsetzung $\emptyset^{\varkappa^{-1}} = \emptyset$. Schließlich definieren wir noch $\varkappa^0$ durch $L^{\varkappa^0} = L$ für $L \in \mathfrak{L}_0$, $L^{\varkappa^0} = \emptyset$ für $L \notin \mathfrak{L}_0$ und $\emptyset^{\varkappa^0} = \emptyset$. Nun sind in üblicher Weise alle ganzzahligen Potenzen von $\varkappa$ erklärt und wir können sofort die *Potenzregel* $\varkappa^{m+n} = \varkappa^m \varkappa^n$ verifizieren.

Wir definieren jetzt durch

$$\varkappa_{ij} = \begin{cases} \varkappa^{j-i} & i \neq j \\ \underline{1} & i = j \end{cases} \qquad (i, j \in I = \{1, 2, 3\}) \tag{4.2}$$

die kanonischen Abbildungen $\varkappa_{ij}$ zu unserer gesuchten Struktur. Wir müssen zeigen, daß sie den Bedingungen (I) bis (VI) genügen. Unmittelbar aus (4.2) folgt (II), zusammen mit (4.1) auch (I) und (IV). Ferner ergibt sich sofort, daß alle $\mathfrak{L}(i, j)$ mit $i \neq j$ übereinstimmen; in diesem Fall gilt nämlich $\mathfrak{L}(i, j) = \mathfrak{L}_0$. Aus (4.2) und der oben zitierten Potenzregel folgt $\varkappa_{ij}\varkappa_{jk} = \varkappa_{ik}$, also auch $L^{\varkappa_{ij}\varkappa_{jk}} = L^{\varkappa_{ik}}$, mithin insbesondere (III).

Wir beweisen jetzt (V), d.h. daß $\mathfrak{P}^{\varkappa^k}$ [$k = j - i$ ($\neq 0$) gesetzt] eine Zerlegung von $\mathfrak{k}$ ist. Zunächst bemerken wir, *daß* $L_1^{\tau_1}$ *und* $L_2^{\tau_2}$ ($\tau_1, \tau_2 \in \mathsf{T}$) *genau dann einen nichtleeren Durchschnitt haben, wenn es ein* $\tau^* \in \mathsf{T}$ *mit*

$$\tau_i \equiv \tau^* \quad (\mathsf{T}(L_i)) \qquad (i = 1, 2) \tag{4.3}$$

gibt. In der Tat, gilt (4.3), so ist $L_1^{\tau_1} \cap L_2^{\tau_2} = L_1^{\tau^*} \cap L_2^{\tau^*} = \{t^{\tau^*}\}$ ($L_1 \cap L_2 = \{t\}$ sei das Zentrum des Büschels $\mathfrak{P}$); existiere nun andererseits ein $u \in T$ mit $\{u\} = L_1^{\tau_1} \cap L_2^{\tau_2}$, es gibt ein $\tau^* \in \mathsf{T}$ mit $u = t^{\tau^*}$, hieraus folgt $L_i^{\tau_i} = L_i^{\tau^*}$ ($i = 1, 2$), mithin (4.3). Nun können wir zeigen: *Die Geraden* $L_1^{\varkappa^k}$ *und* $L_2^{\varkappa^k}$ [$L_1, L_2 \in \mathfrak{P} \cap \mathfrak{L}(i, j)$] *sind windschief.* Jedenfalls sind nämlich $L_1^{\varkappa^k}$ und $L_2^{\varkappa^k}$ nicht zueinander parallel, weil L_1 und L_2 es nicht sind. Wir müssen also jetzt noch $L_1^{\varkappa^k} \cap L_2^{\varkappa^k} = \emptyset$ beweisen. Wir nehmen das Gegenteil an. Aus $L_i^{\varkappa} \parallel L_i$ folgt die Existenz von $\tau_i \in \mathsf{T}$ mit $L_i^{\varkappa} = L_i^{\tau_i}$, hieraus $L_i^{\varkappa^k} = L_i^{\tau_i^k}$. Nach der zu (4.3) gemachten Bemerkung gibt es auf Grund unserer Annahme ein $\tau^* \in \mathsf{T}$ mit $\tau_i^k \equiv \tau^*$ ($\mathsf{T}(L_i)$). Wegen $|k| = |j - i| < 3 < p$ (p ist die Charakteristik von $\mathfrak{k}$ und auch von $\mathfrak{p}$) existiert ein l mit $kl \equiv 1$ (p), hieraus schließt man $\tau_i^{kl} = \tau_i \equiv \tau^{*l}$ ($\mathsf{T}(L_i)$). Wiederum wegen (4.3) würde daraus aber $L_1^{\varkappa} \cap L_2^{\varkappa} = L_1^{\tau_1} \cap L_2^{\tau_2} \neq \emptyset$ folgen, im Widerspruch zu unserer Annahme, daß $\mathfrak{Z} = \mathfrak{P}^{\varkappa}$ eine Zerlegung sei.

Da $\mathfrak{k}$ als endlich vorausgesetzt war, folgt aus Abzählungsgründen nunmehr sofort: $\mathfrak{P}^{\varkappa^k}$ *ist eine Zerlegung von* $\mathfrak{k}$. Denn $\mathfrak{P}^{\varkappa}$ ist voraussetzungsgemäß eine

Zerlegung, d.h. die Geraden von $\mathfrak{P}^x$ überdecken die Punktmenge T von t. Mithin ist $|T| = \sum_{L \in \mathfrak{P}^x} |L|$. Da $\mathfrak{P}^{x^2} = \mathfrak{P}^{x \cdot x^{p-1}}$ aus allen Geraden $L^{x^{p-1}}$ mit $L \in \mathfrak{P}^x$ besteht, $|L^{x^{p-1}}| = |L|$ ist, und die Geraden aus $\mathfrak{P}^{x^{p-1}}$ paarweise windschief liegen, folgt, daß auch sie ganz T überdecken. Damit sind für $\mathfrak{P}^{x^2}$ alle Eigenschaften einer Zerlegung gezeigt.

Es bleibt also nur noch (VI) zu beweisen. Das tun wir, indem wir alle möglichen Fälle durchdiskutieren:

(a) $\varrho = (123)$, $\sigma = \underline{1}$: Es ist $x_{13} \neq x_{21}$ wegen $x^2 \neq x^{-1}$, $x^3 \neq \underline{1}$; denn wir haben $p \geqq 5$ vorausgesetzt.

(b) $\varrho = (123)$, $\sigma \neq \underline{1}$ (wobei wir o.B.d.A. t als Zentrum der Streckung σ annehmen können): Es ist $x_{12} \neq \sigma^{-1} x_{23} \sigma$. Bei der gegenteiligen Annahme wäre nämlich $\mathfrak{Z} = \mathfrak{P}^x = \mathfrak{P}^{x_{12}} = \mathfrak{P}^{\sigma^{-1} x_{12} \sigma} = \mathfrak{P}^{x_{12}\sigma} = \mathfrak{P}^{x\sigma} = \mathfrak{Z}^\sigma$; dies ist aber für $\sigma \neq \underline{1}$ nicht möglich, denn es gibt $L \in \mathfrak{Z}$, die nicht durch t gehen, und für diese ist $L \| L^\sigma \neq L$. Ganz entsprechend sieht man:

(c) $\varrho = (132)$, $\sigma = \underline{1}$: Es ist $x_{12} \neq x_{31}$.

(d) $\varrho = (132)$, $\sigma \neq \underline{1}$: Es ist $x_{12} \neq \sigma^{-1} x_{23} \sigma$.

Damit haben wir alle überhaupt möglichen Fälle durchdiskutiert; (VI) ist erfüllt. *Unsere Struktur $\mathfrak{p} = \{t, x^{j-i}\}$ ist also eine T-Struktur mit genau drei Transitivitätsgebieten hinsichtlich ihrer Translationsgruppe* T.

3. Die einfachsten T-Strukturen dieser Art gewinnt man auf folgende Weise: Sei $\mathfrak{R}$ der dreidimensionale projektive Raum über dem Galois-Feld der Ordnung 5. In $\mathfrak{R}$ sei eine Kongruenz $\mathfrak{K}$ gegeben, d.h. eine Menge paarweise windschiefer Geraden, die $\mathfrak{R}$ überdecken. [Über die Bildung solcher Kongruenzen s. ANDRÉ (1954), insbesondere §8.] Durch Übergang von $\mathfrak{R}$ zum dreidimensionalen affinen Raum $\mathfrak{A}$ werden die eigentlichen Geraden aus $\mathfrak{K}$ zu einer Zerlegung $\bar{\mathfrak{K}} = \mathfrak{Z}$ von $\mathfrak{A}$, wie am Ende von Ziffer 1 näher ausgeführt wurde. Dieses $\mathfrak{Z}$ ist auch Zerlegung einer z.T. $t = [T, \mathfrak{P}]$ (vgl. Ziffer 1)[13]) die $5^3 = 125$ Punkte und $5^2 + 5 + 1 = 31$ Geraden durch jeden Punkt enthält. Nunmehr können wir das in Ziffer 2 gebrachte Konstruktionsverfahren anwenden und erhalten eine T-Struktur mit $3 \cdot 125 = 375$ Punkten, 31 Geraden durch jeden Punkt, $6 \cdot 5 \cdot 5 \cdot 3 = 450$ Geraden mit je 5 Punkten (auf denen die Translationsgruppe transitiv ist) und $25(31-6) = 25^2 = 625$ Geraden mit je $5 \cdot 3 = 15$ Punkten.

§ 5. Über die Nichtexistenz von Parallelstrukturen mit bestimmten Eigenschaften

1. Im vorigen Paragraphen haben wir Beispiele von T-Strukturen angegeben, deren Translationsgruppe nicht transitiv auf den Punkten ist. Wir bringen jetzt zwei Sätze, die zeigen, daß die Anzahl der Transitivitätsgebiete einer T-Struktur nicht beliebig vorgeschrieben werden kann. Untersuchungen dieser Art können für Anwendungen auf affine Ebenen wichtig sein.

[13]) Hier ist T die elementar abelsche Gruppe vom Typ (5, 5, 5) und $\mathfrak{P}$ die Partition der eindimensionalen Untergruppen von T.

SATZ 5.1. *Es gibt keine T-Struktur, deren Punktmenge durch ihre Translationsgruppe in genau zwei Transitivitätsgebiete zerfällt.*

BEWEIS. Sei $\mathfrak{p}=\{t, \varkappa_{ij}\}$ eine T-Struktur mit genau zwei Transitivitätsgebieten hinsichtlich ihrer Translationsgruppe T; die Matrix $(\varkappa_{ij})$ ist dann zweireihig, und wir können $I=\{1, 2\}$ setzen. Wir müssen zeigen, daß wir die $\varkappa_{ij}$ nicht so wählen können, daß die Bedingungen (I) bis (VI) von Satz 1.2 alle erfüllt sind (s. §1, Ziffer 4). Hierzu machen wir zunächst die folgende Vorbemerkung: Sei $t=[T, \mathfrak{P}]$ der Stamm von $\mathfrak{p}=\{t, \varkappa_{ij}\}$. Nach II, §§1, 2 können wir T als eine zu T isomorphe Gruppe (mit dem neutralen Element o) auffassen. Die durch $p^{\varepsilon}=-p$ $(p\in T)$ definierte Abbildung ε ist eine o-Streckung von t, was sofort aus II, Satz 4.2, folgt. Es ist $\varepsilon^2=\underline{1}$, ferner gilt

$$\tau^{-1}=\varepsilon\tau\varepsilon \tag{5.1}$$

für jedes $\tau\in\mathsf{T}$. Aus $o^{\tau}=t$ folgt nämlich $p^{\tau}=p+t$ (vgl. II, §1), und wir erhalten:

$$p^{\tau^{-1}}=p-t=-(-p+t)=(-p+t)^{\varepsilon}=(-p)^{\tau\varepsilon}=p^{\varepsilon\tau\varepsilon}$$

für jedes $p\in T$, also (5.1). Hieraus und aus den Folgerungen (f), (i) und (j) der Bedingungen (I) bis (VI) für kanonische Abbildungen (§1, Ziffer 5) erhalten wir jetzt $(\tau_{12,L}=\tau$ gesetzt)

$$L^{\varkappa_{21}}=L^{\varkappa_{11}^{-1}}=L^{\tau^{-1}}=L^{\varepsilon\tau}=L^{\varepsilon\varkappa_{12}\varepsilon}$$

für alle $L\in\mathfrak{L}$ mit $L^{\varkappa_{11}}\neq\emptyset$. Wegen Folgerung (e) ist aber $L^{\varkappa_{11}}=\emptyset$ gleichbedeutend mit $L^{\varkappa_{21}}=\emptyset$. Wir erhalten also

$$\varkappa_{21}=\varepsilon\varkappa_{12}\varepsilon=\varepsilon^{-1}\varkappa_{12}\varepsilon. \tag{5.2}$$

Dies widerspricht aber der Bedingung (VI) für $\varrho=(12)$ und $\sigma=\varepsilon$. Damit ist unser Satz bewiesen.

2. Jetzt sei $\mathfrak{p}=(P, \mathfrak{G})$ eine *endliche* T-Struktur deren Punktmenge P durch die Gruppe Π aller affin-axialen Kollineationen von $\mathfrak{p}$ in genau zwei Transitivitätsgebiete P_0 und P_1 zerlegt wird. Zur Untersuchung solcher T-Strukturen knüpfen wir an die Bemerkungen in Ziffer 5 von §3 an: O.B.d.A. sei P_0 auch Transitivitätsgebiet hinsichtlich der Translationsgruppe T, d.h. P_0 sei der Sockel von $\mathfrak{p}$. Durch T zerfällt P_1 in s Transitivitätsgebiete T_i $(i\in I)$, hierbei ist s die Ordnung jeder Streckungsgruppe Π_{u_0} mit $u_0\in P_0$. Als Indexmenge nehmen wir $I=\{0\}\cup\Pi_{u_0}$ (für ein festes $u_0\in P_0$), und zwar indizieren wir die Transitivitätsgebiete von T auf folgende Weise: Dem Sockel geben wir den Index 0, d.h. $P_0=T_0$. Einem beliebigen in P_1 gelegenen Transitivitätsgebiet von T geben wir $\underline{1}\in\Pi_{u_0}$ als Index. Jedes andere Transitivitätsgebiet in P_1 läßt sich aus $T_{\underline{1}}$ durch Anwendung genau einer Streckung $\sigma\in\Pi_{u_0}$ gewinnen; diesem geben wir den Index σ, d.h. wir setzen $T_{\sigma}=T_{\underline{1}}^{\sigma}$.

Um nun eine Übersicht über die zu $\mathfrak{p}$ gehörenden kanonischen Abbildungen zu bekommen, wenden wir zunächst Satz 3.4 an, der uns eine Übersicht über die Streckungen von $\mathfrak{p}$ gab. Formel (3.6) ergibt für $j=0$ in unserem Fall wegen der von uns soeben vorgenommenen Indizierung

$$\varkappa_{0\sigma}=\sigma^{-1}\varkappa_{0\underline{1}}\sigma. \tag{5.3}$$

[Hierbei haben wir das in (3.4) vorkommende σ^* mit σ identifiziert; dies konnten wir tun, weil σ durch die Wirkung auf die Punkte von $\underline{1}$, also durch σ^*, eindeutig bestimmt ist.]

Aus (5.3) erhalten wir $\mathfrak{L}(0,\sigma)=\mathfrak{L}(0,\underline{1})$. [Zur Definition von $\mathfrak{L}(i,j)$ vgl. §1, Ziffer 5.] Mit Hilfe von (1.3) und (1.4) können wir nun $\mathfrak{L}(0,\sigma)=\mathfrak{L}(\sigma,0)=\mathfrak{L}(0,\sigma')=\mathfrak{L}(\sigma,0)\cap\mathfrak{L}(0,\sigma')\subseteqq\mathfrak{L}(\sigma,\sigma')$, also auch $\mathfrak{P}(0,\sigma)=\mathfrak{P}\cap\mathfrak{L}(0,\sigma)\subseteqq\mathfrak{P}(\sigma,\sigma')$ schließen. ($\mathfrak{P}$ ist das Büschel aller Geraden durch einen festen Punkt von T.) Sei $\mathfrak{P}(0,\sigma)=\{L^{(0)},\ldots,L^{(m)}\}$ und $\mathfrak{P}(\sigma,\sigma')=\{L^{(0)},\ldots,L^{(m')}\}$; wegen $\mathfrak{P}(0,\sigma)\subseteqq\mathfrak{P}(\sigma,\sigma')$ ist $m\leqq m'$. Nach (V) sind $\mathfrak{P}^{\varkappa_0\sigma}=\mathfrak{P}(0,\sigma)^{\varkappa_0\sigma}$ und für $\sigma\neq\sigma'$ auch $\mathfrak{P}^{\varkappa_{\sigma\sigma'}}$ Zerlegungen von t. Daraus folgt

$$|T|=\sum_{r=0}^{m}|L^{(r)}|=\sum_{r=0}^{m'}|L^{(r)}|$$

und somit $m=m'$, $\mathfrak{P}(0,\sigma)=\mathfrak{P}(\sigma,\sigma')$, also

$$\mathfrak{L}(0,\sigma)=\mathfrak{L}(\sigma,\sigma')\quad\text{für}\quad\sigma\neq\sigma'. \tag{5.4}$$

Sämtliche $\mathfrak{L}(i,j)$ $(i,j\in I=\{0\}\cup\Pi_{\underline{u}_0})$ *mit* $i\neq j$ *stimmen also überein.* Somit vereinfacht sich (III) zu $\varkappa_{ij}\varkappa_{jk}=\varkappa_{ik}$ $(i,j,k\in I)$. Mit (5.3) erhalten wir jetzt

$$\varkappa_{\sigma\sigma'}=\varkappa_{\sigma 0}\varkappa_{0\sigma'}=\varkappa_{0\sigma}^{-1}\varkappa_{0\sigma'}=\sigma^{-1}\varkappa_{0\underline{1}}^{-1}\sigma\sigma'^{-1}\varkappa_{0\underline{1}}\sigma'. \tag{5.5}$$

Durch $\varkappa_{0\underline{1}}$ *sind somit alle kanonischen Abbildungen eindeutig bestimmt.*

3. Wir wollen nun den folgenden, für eine Anwendung auf affine Ebenen nützlichen Satz herleiten:

SATZ 5.2. *Es gibt keine endliche T-Struktur* $\mathfrak{p}=(P,\mathfrak{G})$ *der Charakteristik* p *mit den beiden folgenden Eigenschaften:*

(a) *Durch die Gruppe* Π *aller affin-axialen Kollineationen von* $\mathfrak{p}$ *zerfällt die Menge* P *der Punkte von* $\mathfrak{p}$ *in genau zwei Transitivitätsgebiete.*

(b) *Die Anzahl der Transitivitätsgebiete, in die* P *durch die Translationsgruppe* T *von* $\mathfrak{p}$ *zerlegt wird, ist eine Potenz der Charakteristik* p *von* $\mathfrak{p}$.

Dem Beweis schicken wir einige Hilfssätze voraus.

HILFSSATZ 5.1[14]). *Die Streckungsgruppe* $\Pi_{\underline{u}_0}$ *ist (bis auf Isomorphie) die multiplikative Gruppe* $\mathsf{K}_0^\times$ *eines Galois-Feldes* K_0, *das Unterkörper des Kernes* K (vgl. II, §4) *des Stammes* t *von* $\mathfrak{p}$ *ist.*

BEWEIS. Nach Voraussetzung (b) von Satz 5.2 gibt es eine natürliche Zahl b derart, daß p^b die Anzahl der Transitivitätsgebiete hinsichtlich T ist, also gilt $|\Pi_{\underline{u}_0}|=p^b-1$. Die Ordnung von K sei p^a. Aus $p^b-1\,|\,p^a-1=|\mathsf{K}^\times|$ folgt [vgl. ANDRÉ (1958), Hilfssatz 3.6] $b\,|\,a$, es gibt mithin einen Unterkörper K_0 von K mit der Ordnung p^b. Da $\mathsf{K}^\times$ als multiplikative Gruppe eines Galois-Feldes zyklisch ist, enthält $\mathsf{K}^\times$ nur eine Untergruppe der Ordnung p^b-1, also ist $\mathsf{K}_0^\times=\Pi_{\underline{u}_0}$.

Wir können jetzt die 0 unserer Indexmenge I als Nullelement von K_0 auffassen; tun wir dies, so wird $I=\mathsf{K}_0$.

[14]) Vgl. auch ANDRÉ (1958), Hilfssatz 3.7.

HILFSSATZ 5.2. *Für alle* $\alpha \in \mathsf{K}_0$ *gilt*

$$\varkappa_{\underline{1},\alpha+\underline{1}} = \varkappa_{0\alpha}. \tag{5.6}$$

Zum Beweis machen wir die folgende Fallunterscheidung:

(a) $\alpha = 0$: Nach (II) ist in diesem Falle $\varkappa_{\underline{1}\underline{1}} = \varkappa_{00} = \underline{1}$.

(b) $\alpha + \underline{1} = 0$, d.h. $\alpha = -\underline{1} = \varepsilon$: Aus (5.3) und (5.2)[15]) erhalten wir $\varkappa_{0\varepsilon} = \varepsilon^{-1}\varkappa_{0\underline{1}}\varepsilon = \varkappa_{\underline{1}0}$, das ist aber (5.6) für unseren Fall.

(c) $\alpha \neq 0$, ε (d.h. α^{-1} und $(\alpha+\underline{1})^{-1}$ existieren): Um (5.6) für diesen Fall zu zeigen, verifizieren wir zunächst die folgende Identität:

$$\tau(\alpha+\underline{1}) = (\alpha+\underline{1})\,\tau\alpha^{-1}\,\tau\alpha; \tag{5.7}$$

hierbei ist τ eine beliebige Translation. Sei nämlich $q \in T$ ein beliebiger Punkt des Stammes t von $\mathfrak{p}$. Wir können $q^\tau = q + t$ für ein geeignetes $t \in T$ setzen (vgl. II, §1) und erhalten

$$\begin{aligned} q^{\tau(\alpha+\underline{1})} &= (q+t)^{\alpha+\underline{1}} = q^\alpha + q + t^\alpha + t = (q + q^{\alpha^{-1}} + t + t^{\alpha^{-1}})^\alpha = (q + q^{\alpha^{-1}} + t^{\alpha^{-1}})^{\tau\alpha} \\ &= (q^\alpha + q + t)^{\alpha^{-1}\tau\alpha} = (q^\alpha + q)^{\tau\alpha^{-1}\tau\alpha} = q^{(\alpha+\underline{1})\tau\alpha^{-1}\tau\alpha}, \end{aligned}$$

mithin (5.7). Man gewinnt daraus nach einfacher Umrechnung

$$\tau^{-1}(\alpha+\underline{1})^{-1}\,\tau(\alpha+\underline{1}) = \alpha^{-1}\,\tau\alpha. \tag{5.8}$$

Sei nun $L \in \mathfrak{L}(i,j)$ $[i \neq j]$. Nach Folgerung (i) [vgl. §1, Ziffer 5] gibt es eine Translation $\tau_{0\underline{1},L} = \tau$ mit $L^{\varkappa_{0\underline{1}}} = L^\tau$. Mit (5.3), (5.8) und (5.5) erhalten wir jetzt

$$L^{\varkappa_{0\alpha}} = L^{\overset{\varepsilon}{\alpha^{-1}\varkappa_{0\underline{1}}\alpha}} = L^{\alpha^{-1}\tau\alpha} = L^{\tau^{-1}(\alpha+\underline{1})^{-1}\tau(\alpha+\underline{1})} = L^{\varkappa_{\underline{1}}(\alpha+\underline{1})^{-1}\varkappa_{0\underline{1}}(\alpha+\underline{1})} = L^{\varkappa_{\underline{1}0}\varkappa_{0,\alpha+\underline{1}}} = L^{\varkappa_{\underline{1},\alpha+\underline{1}}}$$

für alle $L \in \mathfrak{L}(i,j)$. Ist $L \notin \mathfrak{L}(i,j)$, so ist $L^{\varkappa_{0\alpha}} = L^{\varkappa_{\underline{1},\alpha+\underline{1}}} = \emptyset$. Damit ist (5.6) vollständig gezeigt.

HILFSSATZ 5.3. *Für alle* $\alpha, \beta \in \mathsf{K}_0$ *gilt*

$$\varkappa_{\alpha+\underline{1},\beta+\underline{1}} = \varkappa_{\alpha\beta}. \tag{5.9}$$

Dies folgt sofort aus Hilfssatz 5.2 und (III) durch

$$\varkappa_{\alpha+\underline{1},\beta+\underline{1}} = \varkappa_{\alpha+\underline{1},\underline{1}}\varkappa_{\underline{1},\beta+\underline{1}} = \varkappa^{-1}_{\underline{1},\alpha+\underline{1}}\varkappa_{\underline{1},\beta+\underline{1}} = \varkappa^{-1}_{0\alpha}\varkappa_{0\beta} = \varkappa_{\alpha 0}\varkappa_{0\beta} = \varkappa_{\alpha\beta}.$$

Nunmehr sind wir in der Lage, Satz 5.2 zu beweisen. Die Abbildung $\alpha \to \alpha + \underline{1}$ ist nämlich eine auf $I = \mathsf{K}_0$ reguläre Permutation, und (5.9) ergibt somit einen Widerspruch zu Bedingung (VI) für kanonische Abbildungen. Eine in Satz 5.2 angegebene T-Struktur kann es also nicht geben.

4. Wir wenden jetzt Satz 5.2 auf endliche affine Ebenen an und beweisen einen Satz, der in ANDRÉ (1958) bereits auf etwas anderem Wege gezeigt wurde. Unter einer *T-Ebene* verstehen wir eine affine Ebene, die zugleich T-Struktur ist, m.a.W. eine affine Ebene, in der es nichtidentische Translationen in jeder Richtung gibt.

[15]) Die Formel (5.2) gilt nämlich, wie aus ihrer Herleitung sofort ersichtlich, allgemein und nicht nur für den dort behandelten Fall zweier Transitivitätsgebiete.

SATZ 5.3. *Es gibt keine endliche T-Ebene mit genau zwei Transitivitätsgebieten hinsichtlich ihrer Gruppe der affin-axialen Kollineationen (d.h. Translationen und Streckungen)*[18].

BEWEIS. Sei $\mathfrak{A}$ eine endliche T-Ebene mit der im Satz genannten Eigenschaft. Nach Hilfssatz 3.5 von ANDRÉ (1958) ist die Anzahl der Transitivitätsgebiete hinsichtlich der Translationsgruppe T von $\mathfrak{A}$ eine Potenz der Charakteristik p. Da $\mathfrak{A}$ voraussetzungsgemäß eine T-Struktur ist, folgt die Nichtexistenz jetzt sofort aus Satz 5.2.

[18]) In ANDRÉ (1958) wurde, etwas allgemeiner als hier, sogar die Nichtexistenz von Ebenen mit nichtidentischen Translationen in mindestens zwei Richtungen (sog. t-Ebenen) mit diesen Eigenschaften bewiesen. Dieser Fall läßt sich aber leicht auf den von Satz 5.3 zurückführen, wie dort gezeigt wurde.

Literatur

ANDRÉ, J.: Über nicht-desarguessche Ebenen mit transitiver Translationsgruppe. Math. Z. **60**, 156–186 (1954).

— Über Perspektivitäten in endlichen affinen Ebenen. Arch. Math. **9**, 228–235 (1958).

— Affine Ebenen mit genügend vielen Translationen. Math. Nachr. **19**, 204–210 (1959).

ZASSENHAUS, H.: Lehrbuch der Gruppentheorie, Bd. I. Leipzig: Teubner 1937.

Braunschweig, Mathematisches Institut der Technischen Hochschule

(Eingegangen am 24. Oktober 1960)

LEICHTWEISS, K.
Math. Zeitschr. 76, 334–366 (1961)

Zur Riemannschen Geometrie in Grassmannschen Mannigfaltigkeiten

Von

KURT LEICHTWEISS

Bekanntlich kann jede (reelle) differenzierbare Mannigfaltigkeit M mit (mindestens) einer Riemannschen Metrik versehen werden[1]). Um nun eine Auswahl unter allen möglichen Riemannschen Metriken von M zu treffen, kann man eine derartige Metrik etwa dadurch auszeichnen, daß man die Invarianz derselben gegenüber einer vorgegebenen Gruppe G von homöomorphen und differenzierbaren Selbstabbildungen von M verlangt. Besonders einfach werden die Verhältnisse, wenn M schon als homogener Raum G/g vorausgesetzt wird (G = zusammenhängende Liesche Gruppe, g = abgeschlossene Untergruppe von G), da dann G in einer einfach zu beschreibenden Weise auf M operiert. Es ist leicht einzusehen, daß in $M = G/g$ sicher eine gegen G invariante Riemannsche Metrik existiert, wenn g kompakt ist. É. CARTAN konnte sogar zeigen, daß ein sog. „symmetrischer" homogener Raum M im Falle einer einfachen Gruppe G auch (von einem positiven konstanten Maßfaktor abgesehen) nur eine gegen G invariante Riemannsche Metrik besitzt[2]). Außerdem wies CARTAN u.a. nach, daß die Riemannsche Krümmung aller gegen G invarianten Riemannschen Metriken eines symmetrischen und kompakten M überall nichtnegativ ist.

Musterbeispiele von kompakten und symmetrischen homogenen Räumen werden durch die reellen Grassmannschen Mannigfaltigkeiten

$$G_{l,m} = SO(l+m)/\big(O(l)\times O(m)\big) \cap SO(l+m)$$

und

$$\widetilde{G_{l,m}} = SO(l+m)/SO(l)\times SO(m)$$

bzw. die komplexen Grassmannschen Mannigfaltigkeiten

$$H_{l,m} = SU(l+m)/\big(U(l)\times U(m)\big) \cap SU(l+m)$$

($l \geqq 1$, $m \geqq 1$; $O(n)$ bzw. $SO(n)$ bzw. $U(n)$ bzw. $SU(n)$ = orthogonale bzw. spezielle orthogonale bzw. unitäre bzw. spezielle unitäre Gruppe vom Grade n) geliefert. Wegen der Einfachheit von $SO(n)$ im Falle $n \neq 4$ bzw. von $SU(n)$ allgemein muß also nach CARTAN in $G_{l,m}$ und $\widetilde{G_{l,m}}$ im Falle $l+m \neq 4$ bzw. in $H_{l,m}$ allgemein eine durch Invarianz gegen $SO(l+m)$ bzw. $SU(l+m)$ ausgezeichnete, „kanonische" Riemannsche Metrik mit interessanten Eigen-

[1]) Siehe [1], theorem 4.

[2]) In Abschnitt 4.1 wird näher auf die Resultate CARTANS eingegangen.

schaften existieren, während der Fall $l+m=4$ bei $G_{l,m}$ und $\widetilde{G}_{l,m}$ offenbar eine Ausnahme darstellt. Diese Tatsachen berechtigen uns zu der folgenden Frage: *Wie sieht eigentlich die kanonische Metrik der angegebenen Grassmannschen Mannigfaltigkeiten aus, und welche invarianten Riemannschen Metriken existieren im Ausnahmefall $l+m=4$ in $G_{l,m}$ und $\widetilde{G}_{l,m}$?* Die Antwort auf diese Frage bildet das Hauptanliegen unserer Arbeit.

Es erweist sich nun für die diesbezüglichen Überlegungen als zweckmäßig, nicht die von CARTAN verwandte und auf der Untersuchung von Lieschen Gruppen beruhende Methode heranzuziehen, sondern einen mehr geometrischen Weg einzuschlagen und von der Definition der Grassmannschen Mannigfaltigkeiten als Menge aller l-dimensionalen Untervektorräume des $(l+m)$-dimensionalen reellen bzw. komplexen Vektorraums auszugehen. Geeignete n-Beine von Vektoren als Funktionen von ebenfalls geeigneten lokalen Koordinaten der Grassmannschen Mannigfaltigkeiten liefern dann sowohl für die kanonische Metrik als auch für sämtliche gegen $SO(l+m)$ invarianten Riemannschen Metriken von $G_{l,m}$ und $\widetilde{G}_{l,m}$ im Fall $l+m=4$ brauchbare explizite Ausdrücke (Satz 1, Satz 3 und Satz 6). (In diesen treten innere Produkte bzw. Hermitesche Bilinearformen auf, die auf die Abhängigkeit der gefundenen Metriken von den Gruppen $SO(l+m)$ bzw. $SU(l+m)$ hinweisen!) Dabei zeigt sich, *daß in Wirklichkeit nur $l=m=2$ (bei reellen Grassmannschen Mannigfaltigkeiten) einen Ausnahmefall darstellt, dessen Sonderstellung auch durch die bekannte Tatsache $\widetilde{G}_{2,2}=S^2\times S^2$ ($S^2=$ zweidimensionale Sphäre) erklärbar ist. Hier hängt die Gesamtheit aller invarianten Riemannschen Metriken von zwei Parametern ab, während sie sonst stets einparametrig ist.* Der (nach einigen Vorbereitungen in Abschnitt 1.) in Abschnitt 2. dargelegte Beweis für diese Hauptresultate gelingt durch Zurückführung des Problems der Bestimmung aller gegen G invarianten Riemannschen Metriken auf das Problem, gewisse Darstellungen von g im Tangentialraum eines Punktes der betreffenden Grassmannschen Mannigfaltigkeit in irreduzible Bestandteile zu zerlegen, wozu nur wenige Tatsachen der Darstellungstheorie benötigt werden.

Im dritten Abschnitt der Arbeit werden weiter die Krümmungseigenschaften der Grassmannschen Mannigfaltigkeiten in bezug auf die invarianten Riemannschen Metriken untersucht. Mittels des Cartanschen Differentialformenkalküls lassen sich Krümmungstensor und somit auch Ricci-Tensor, Skalarkrümmung und Riemannsche Krümmung leicht explizit bestimmen. Man erhält dadurch die folgenden, teilweise Ergebnisse CARTANS bestätigenden Resultate: *Die mit der (einparametrigen) Riemannschen „Normalmetrik" versehenen Grassmannschen Mannigfaltigkeiten sind Einsteinsche Räume im Gegensatz zu den mit der (zweiparametrigen) „Ausnahmemetrik" versehenen Mannigfaltigkeiten $G_{2,2}$ und $\widetilde{G}_{2,2}$. Die Riemannsche Krümmung ist in jedem Fall überall nichtnegativ, aber nur dann überall positiv, wenn $l=1$ oder $m=1$ gilt, d.h. wenn es sich um (reelle bzw. komplexe) projektive Räume mit elliptischer Metrik handelt* (Satz 4, Satz 5 und 4.2).

In Abschnitt 4 fogt schließlich unter anderem die Charakterisierung der „Normalmetrik" der als algebraische Untermannigfaltigkeiten von projektiven Räumen aufgefaßten Grassmannschen Mannigfaltigkeiten als diejenige Riemannsche Metrik, die durch die elliptische Metrik der projektiven Räume induziert wird (Satz 7). Außerdem werden anstelle von invarianten Riemannschen Metriken allgemeiner invariante nichtausgeartete quadratische bzw. Hermitesche Differentialformen in unsere Betrachtungen miteinbezogen sowie „inhomogene" Grassmannsche Mannigfaltigkeiten $I_{l,m}$ untersucht, welche als Menge aller l-dimensionalen linearen Unterräume des $(l+m)$-dimensionalen reellen affinen Raums definiert sind. Dabei zeigt sich, daß $I_{l,m}$ zwar keine gegenüber der speziellen Bewegungsgruppe $SB(l+m)$ invariante Riemannsche Metrik (Satz 8), wohl aber ein gegen $SB(l+m)$ invariantes Volumelement besitzt, welches in der Integralgeometrie eine Rolle spielt.

Es sei noch darauf hingewiesen, daß C. TELEMAN in einer kürzlich erschienenen Arbeit[3]) ähnliche Untersuchungen über Grassmannsche Mannigfaltigkeiten — allerdings mit anderen Methoden — angestellt hat.

1. Problemstellung und Ergebnisse über invariante Riemannsche Metriken in Grassmannschen Mannigfaltigkeiten

Bekanntlich ist die reelle Grassmannsche Mannigfaltigkeit $G_{l,m}$ als Menge aller $(l-1)$-dimensionalen projektiven Unterräume des $(l+m-1)$-dimensionalen reellen projektiven Raums — oder, was gleichbedeutend ist, — als Menge aller (unorientierten) l-dimensionalen Untervektorräume des $(l+m)$-dimensionalen reellen Vektorraums definiert ($l \geqq 1$, $m \geqq 1$). Ein „Punkt" o von $G_{l,m}$ wird dann etwa durch ein l-Bein von linear unabhängigen Vektoren $\mathring{e}_i$ des n-dimensionalen Vektorraums V_n dargestellt ($n = m+l$; $i = 1, \ldots, l$), welche den betreffenden Untervektorraum aufspannen. Ergänzt man dies l-Bein $\{\mathring{e}_i\}$ durch das m-Bein $\{\mathring{e}_\alpha\}$ ($\alpha = l+1, \ldots, n$) zu einem aus n linear unabhängigen Vektoren bestehenden n-Bein, so lassen sich die Untervektorräume der Nachbarpunkte von o durch Vektoren der Gestalt

$$\mathfrak{f}_i = \mathring{e}_i + \sum_{\alpha=l+1}^{n} u_{i\alpha}\,\mathring{e}_\alpha \qquad (i = 1, \ldots, l) \tag{1}$$

aufspannen. Zu verschiedenen Koeffizientengesamtheiten $(u_{i\alpha})$ von je lm reellen Zahlen gehören dabei verschiedene Untervektorräume, d.h. also auch verschiedene Punkte von $G_{l,m}$. Die Grassmannsche Mannigfaltigkeit ist infolgedessen eine reell analytische Mannigfaltigkeit der Dimension lm, und die Größen $u_{i\alpha}$ ($i = 1, \ldots, l$; $\alpha = l+1, \ldots, n$) bilden die durch das n-Bein $\{\mathring{e}_i, \mathring{e}_\alpha\}$ festgelegten Koordinaten der Umgebung des Punktes $o \in G_{l,m}$.

Ziel dieses Abschnittes ist es nun, auf der differenzierbaren Mannigfaltigkeit $G_{l,m}$ zunächst eine Riemannsche Metrik zu definieren, welche invariant ist gegenüber allen „eigentlich orthogonalen" Transformationen des V_n, d.h. allen linearen, orientierungserhaltenden Abbildungen des V_n, welche eine (ausgezeichnete) symmetrische Bilinearform $\langle \mathfrak{x}\, \mathfrak{y} \rangle$ ($\mathfrak{x}, \mathfrak{y} \in V_n$) mit positiv definiter

[3]) Siehe [2].

quadratischer Form $\langle \mathfrak{x}\, \mathfrak{x}\rangle$ ungeändert lassen. Dazu ist folgendes zu bemerken: Die Elemente der reellen eigentlich orthogonalen Gruppe vom Grade n, welche mit $SO(n)$ bezeichnet sei, bilden jeden l-dimensionalen Untervektorraum des V_n wieder auf einen ebensolchen ab, sie permutieren also die Punkte von $G_{l,m}$ und zwar in transitiver Weise. Die gesuchte Riemannsche Metrik soll nun — präziser formuliert — gegenüber dieser Permutationsgruppe von $G_{l,m}$ invariant sein. Es erweist sich jetzt als zweckmäßig, die das lokale Koordinatensystem von $G_{l,m}$ in o bestimmenden Basisvektoren $\mathring{\mathfrak{e}}_i$ und $\mathring{\mathfrak{e}}_\alpha$ des V_n als orthonormiert vorauszusetzen: $\langle \mathring{\mathfrak{e}}_i\, \mathring{\mathfrak{e}}_j\rangle = \delta_{ij}$, $\langle \mathring{\mathfrak{e}}_\alpha\, \mathring{\mathfrak{e}}_\beta\rangle = \delta_{\alpha\beta}$, $\langle \mathring{\mathfrak{e}}_i\, \mathring{\mathfrak{e}}_\alpha\rangle = 0$ $(1 \leqq i, j \leqq l;\ +1 \leqq \alpha, \beta \leqq n)$. Dann bilden alle den Punkt o festlassende Elemente von $SO(n)$ eine Untergruppe H, die sich aus den durch

$$(2) \qquad \mathring{\mathfrak{e}}'_i = \sum_{j=1}^{l} a_{ji}\, \mathring{\mathfrak{e}}_j, \qquad \mathring{\mathfrak{e}}'_\alpha = \sum_{\beta=l+1}^{n} b_{\beta\alpha}\, \mathring{\mathfrak{e}}_\beta \qquad (1 \leqq i \leqq l;\ l+1 \leqq \alpha \leqq n)$$

gegebenen Transformationen zusammensetzt, wobei (a_{ij}) und $(b_{\alpha\beta})$ orthogonale Matrizen sind, deren Determinanten $\|a_{ij}\|$ und $\|b_{\alpha\beta}\|$ gleiches Vorzeichen haben. Das durch (2) gegebene Element von H führt den durch (1) bestimmten Punkt p von $G_{l,m}$ in einen Punkt p' über, dessen zugehöriger Untervektorraum nach (2) durch die Vektoren

$$\mathfrak{f}'_i = \sum_{j=1}^{l} a_{ji}\, \mathring{\mathfrak{e}}_j + \sum_{\alpha=l+1}^{n} u_{i\alpha} \Big(\sum_{\beta=l+1}^{n} b_{\beta\alpha}\, \mathring{\mathfrak{e}}_\beta \Big)$$

oder ebenso durch die Vektoren

$$(3) \qquad \mathfrak{g}'_i = \sum_{j=1}^{l} a_{ij}\, \mathfrak{f}'_j = \mathring{\mathfrak{e}}_i + \sum_{j=1}^{l} \sum_{\alpha=l+1}^{n} \sum_{\beta=l+1}^{n} a_{ij}\, u_{j\alpha}\, b_{\beta\alpha}\, \mathring{\mathfrak{e}}_\beta \qquad (i = 1, \ldots, l)$$

aufgespannt wird. Vergleich von (3) mit der Darstellung (1) ergibt für die Koordinaten $u'_{i\alpha}$ von p' die Beziehungen:

$$u'_{i\alpha} = \sum_{j=1}^{l} \sum_{\beta=l+1}^{n} a_{ij}\, b_{\alpha\beta}\, u_{j\beta} \qquad (i = 1, \ldots, l;\ \alpha = l+1, \ldots, n).$$

Die Untergruppe H von $SO(n)$ induziert in dem Tangentialraum T von $G_{l,m}$ im Punkt o eine Gruppe Γ von linearen Transformationen:

$$(4) \qquad x'_{i\alpha} = \sum_{j=1}^{l} \sum_{\beta=l+1}^{n} a_{ij}\, b_{\alpha\beta}\, x_{j\beta} \qquad (i = 1, \ldots, l;\ \alpha = 1+l, \ldots, n),$$

die sog. „lineare Isotropiegruppe" von $G_{l,m}$ bezüglich $SO(n)$ ($x_{i\alpha}$ = Komponenten der Vektoren von T!), wobei jedes Element von Γ auf Grund von (4) als Tensorprodukt der beiden durch (2) gegebenen orthogonalen Transformationen aufgefaßt werden kann[4]). Die Gruppe Γ soll deshalb auch das „eingeschränkte[5]) Tensorprodukt" der orthogonalen Gruppen $O(l)$ und $O(m)$ vom Grade l und m genannt werden, welches wir mit $(O(l) \otimes O(m))^*$ bezeichnen wollen. Aus den Orthogonalitätsrelationen der Matrizen (a_{ij}) und $(b_{\alpha\beta})$ folgt,

[4]) Zur Definition des Tensorprodukts von linearen Transformationen vgl. [3], § 1 n^0 4.

[5]) Die Matrizen (a_{ij}) und $(b_{\alpha\beta})$ sind nämlich der Einschränkung $\|a_{ij}\|\, \|b_{\alpha\beta}\| = 1$ unterworfen.

daß die Elemente von $(O(l)\otimes O(m))^*$ orthogonale Transformationen sind, was durch die Formel

$$(O(l)\otimes O(m))^* \subseteq O(l\,m) \tag{5}$$

zum Ausdruck gebracht werden soll.

Nimmt man jetzt einmal an, es existiere auf $G_{l,m}$ eine gegenüber $SO(n)$ invariante Riemannsche Metrik, so ist dies a fortiori auch in bezug auf H der Fall. Die Koeffizienten des metrischen Fundamentaltensors von $G_{l,m}$ im Punkt o sind dann die Koeffizienten einer gegenüber allen Transformationen (4) invarianten, positiv definiten quadratischen Form des Tangentialraums T. Wenn wir aber umgekehrt eine derartige, auf T definierte quadratische Form gefunden haben, so werde die zugehörige Differentialform $(ds^2)_0$ als metrische Fundamentalform von $G_{l,m}$ im Punkt o definiert. Die metrische Fundamentalform in einem beliebigen anderen Punkt q von $G_{l,m}$ sei weiter das Bild von $(ds^2)_0$ bei derjenigen linearen Abbildung von T auf den Tangentialraum von $G_{l,m}$ in q, welche durch irgendein Element von $SO(n)$ induziert wird, das o in q überführt[6]). Auf diese Weise wird $G_{l,m}$ zu einem Riemannschen Raum mit einer Metrik, welche gegenüber $SO(n)$ invariant ist. *Die Existenz einer gegenüber $SO(n)$ invarianten Riemannschen Metrik in $G_{l,m}$ ist also gleichbedeutend mit der Existenz einer gegenüber $\Gamma=(O(l)\otimes O(m))^*$ invarianten, positiv definiten quadratischen Form auf T.*

Wegen (5) läßt sich nun sofort eine quadratische Form angeben, die gegenüber den Transformationen von $(O(l)\otimes O(m))^*$ invariant ist, nämlich $\sum_{i=1}^{l}\sum_{\alpha=l+1}^{n}(x_{i\alpha})^2$. Ihr entspricht nach unseren eben angestellten Überlegungen eine Riemannsche Metrik von $G_{l,m}$, die symbolisch[7]) in der Form

$$ds^2=\sum_{i=1}^{l}\sum_{\alpha=l+1}^{n}\left(\langle e_\alpha\, de_i\rangle\right)^2 \tag{6}$$

geschrieben werden kann und sich tatsächlich als gegenüber $SO(n)$ invariant erweist. Es fragt sich jetzt, ob es noch eine andere gegenüber $SO(n)$ invariante Riemannsche Metrik geben kann. Darüber gibt der folgende Satz Auskunft:

SATZ 1. *Falls nicht gerade $l=m=2$ gilt, ist die durch* (6) *gegebene Riemannsche Metrik, von einem konstanten positiven Proportionalitätsfaktor λ abgesehen, die einzige gegenüber allen Transformationender Gruppe $SO(n)$ invariante Riemannsche Metrik der reellen Grassmannschen Mannigfaltigkeit $G_{l,m}$.*

Im Falle $l=m=2$ wird dagegen die Gesamtheit aller gegen $SO(n)$ invarianten Riemannschen Metriken von $G_{l,m}$ durch den Ausdruck:

$$\begin{cases} ds^2=\lambda\left(\langle e_3\,de_1\rangle^2+\langle e_3\,de_2\rangle^2+\langle e_4\,de_1\rangle^2+\langle e_4\,de_2\rangle^2\right)+\\ \quad+2\mu\left(\langle e_3\,de_1\rangle\langle e_4\,de_2\rangle-\langle e_3\,de_2\rangle\langle e_4\,de_1\rangle\right)\\ \qquad(\lambda,\mu=\text{reell und konstant};\ \lambda>|\mu|) \end{cases} \tag{7}$$

gegeben.

[6]) Wegen der erwähnten Invarianzeigenschaft der auf T definierten quadratischen Form spielt die Auswahl des o in q abbildenden Elements aus $SO(n)$ hier keine Rolle.

[7]) In Abschnitt 3 dieser Arbeit wird diese Symbolik näher erläutert werden.

Man kann nun analog zum bisherigen auch gegenüber der Gruppe aller orthogonalen Transformationen des V_n invariante Riemannsche Metriken von $G_{l,m}$ untersuchen und findet darüber

SATZ 2. *Die durch* (6) *gegebene Riemannsche Metrik ist bis auf einen konstanten positiven Proportionalitätsfaktor* λ *die einzige gegen* $O(n)$ *invariante Riemannsche Metrik auf der reellen Grassmannschen Mannigfaltigkeit* $G_{l,m}$[8]. Hier spielt der Fall $l=m=2$ also keine Ausnahmerolle.

Schließlich stellt sich noch die Frage nach gegenüber $SO(n)$ bzw. $O(n)$ invarianten Riemannschen Metriken auf der Mannigfaltigkeit $\widetilde{G}_{l,m}$ aller *orientierten* l-dimensionalen Untervektorräume des $(l+m)$-dimensionalen reellen Vektorraums. Jetzt gilt:

SATZ 3. *Die Sätze* 1 *und* 2 *bleiben wörtlich richtig, wenn darin* $G_{l,m}$ *überall durch* $\widetilde{G}_{l,m}$ *ersetzt wird.*

2. Beweise der in Abschnitt 1. angeführten Sätze

Wir haben bereits in 1. nachgewiesen, daß den gegenüber $SO(n)$ invarianten Riemannschen Metriken von $G_{l,m}$ umkehrbar eindeutig solche positiv definite quadratische Formen auf einem Tangentialraum T von $G_{l,m}$ entsprechen, welche gegenüber allen Transformationen der linearen Isotropiegruppe von $G_{l,m}$ bezüglich $SO(n)$ invariant sind. Auf dieselbe Weise sieht man ein, daß den gegenüber $O(n)$ invarianten Riemannschen Metriken auf $G_{l,m}$ die gegenüber der linearen Isotropiegruppe von $G_{l,m}$ bezüglich $O(n)$, d.h. gegenüber $O(l)\otimes O(m)$ invarianten positiv definiten quadratischen Formen auf T entsprechen, während den gegen $SO(n)$ bzw. $O(n)$ invarianten Riemannschen Metriken von $\widetilde{G}_{l,m}$ die gegenüber den linearen Isotropiegruppen $SO(l)\otimes SO(m)$ bzw. $SO(l)\otimes O(m)$ invarianten positiv definiten quadratischen Formen auf T eineindeutig zugeordnet sind[9]).

Unsere Aufgabe ist also jetzt, alle gegenüber einer bekannten Gruppe von — wie sich jeweils zeigt — orthogonalen Transformationen invarianten, positiv definiten quadratischen Formen auf T zu bestimmen. Dieses Problem läßt sich auf ein Problem der Darstellungstheorie von Gruppen zurückführen. Dazu benötigen wir:

DEFINITION 1. *Ein Untervektorraum* W *des (reellen) Vektorraums* V *heißt gegenüber einer Gruppe* $\mathfrak{G}$ *von linearen Transformationen von* V *invariant, wenn jedes Element von* $\mathfrak{G}$ *jeden Vektor aus* W *wieder in einem Vektor aus* W *überführt.*

DEFINITION 2. *Ein gegenüber* $\mathfrak{G}$ *invarianter nicht nulldimensionaler Untervektorraum* W *von* V *heißt bezüglich* $\mathfrak{G}$ *reduzibel, wenn* W *einen vom Nullraum*

[8]) Satz 2 enthält für $l=1$ die Tatsache, daß die bis auf einen Faktor einzige, gegen alle Bewegungen des auf dem projektiven Raum $G_{1,m}$ definierten elliptischen Raums E_m invariante Riemannsche Metrik des $G_{1,m}$, die sog. „elliptische" Metrik von $G_{1,m}$, von der Cayley-Kleinschen Metrik des E_m induziert wird.

[9]) Die Tensorprodukte sind hier in Komplexschreibweise zu verstehen.

und W verschiedenen, gegenüber $\mathfrak{G}$ invarianten Untervektorraum W' enthält. Anderenfalls wird W als bezüglich $\mathfrak{G}$ irreduzibel bezeichnet.

Ist speziell V ein euklidischer Vektorraum von endlicher Dimension und $\mathfrak{G}$ eine Gruppe von orthogonalen Transformationen, so ist mit jedem gegen $\mathfrak{G}$ invarianten Untervektorraum W von V auch das orthogonale Komplement von W gegen $\mathfrak{G}$ invariant; man kann also in diesem Falle V als direkte Summe von endlich vielen, paarweise orthogonalen, (bezüglich $\mathfrak{G}$) invarianten und irreduziblen Untervektorräumen darstellen. Jetzt gilt:

HILFSSATZ 1. *Läßt der endlichdimensionale euklidische Vektorraum V (von der Reihenfolge der Summanden abgesehen) nur eine einzige Darstellung als direkte Summe von paarweise orthogonalen und bezüglich der Gruppe $\mathfrak{G}$ von orthogonalen Transformationen von V invarianten sowie irreduziblen Untervektorräumen $W^{(\nu)}$ $(\nu=1,\ldots,s)$ zu, so wird die Gesamtheit aller auf V definierten und gegen $\mathfrak{G}$ invarianten, positiv definiten quadratischen Formen durch*

$$(8)\qquad \sum_{\nu=1}^{s}\lambda_\nu\Big(\sum_{t=n_{\nu-1}+1}^{n_\nu}(y_t)^2\Big)\qquad(\lambda_\nu>0;\ 0=n_0<n_1<\cdots<n_s=\operatorname{Dim}V)$$

gegeben, wenn man für V eine solche orthonormierte Basis zugrundelegt, daß $W^{(\nu)}$ durch $y_t=0$ mit $1\leqq t\leqq n_{\nu-1}$, $n_\nu+1\leqq t\leqq n_s$ dargestellt wird.

BEWEIS. Trivialerweise sind alle quadratischen Formen (8) positiv definit und gegen $\mathfrak{G}$ invariant. Wir nehmen nun an, die (in Matrizenschreibweise dargestellte) quadratische Form $\mathfrak{x}^*\mathfrak{A}\mathfrak{x}$ sei auf V definiert, positiv definit und gegenüber allen orthogonalen Transformationen $\mathfrak{x}'=\mathfrak{T}\mathfrak{x}$ aus $\mathfrak{G}$ invariant ($\mathfrak{A}^*=\mathfrak{A}$, $\mathfrak{T}^*=\mathfrak{T}^{-1}$). Dann folgt aus der Identität $\mathfrak{x}^*\mathfrak{A}\mathfrak{x}\equiv\mathfrak{x}^*(\mathfrak{T}^{-1})^*\mathfrak{A}\mathfrak{T}^{-1}\mathfrak{x}$ zusammen mit $\mathfrak{T}^*=\mathfrak{T}^{-1}$: $\mathfrak{A}=\mathfrak{T}\mathfrak{A}\mathfrak{T}^{-1}$. Wenn jetzt λ ein (sicher reeller) Eigenwert und V_λ der zugehörige Eigenraum der symmetrischen Matrix $\mathfrak{A}$ ist, so gilt für einen beliebigen Vektor $\mathfrak{v}\in V_\lambda$ wegen der eben abgeleiteten Matrixgleichung:

$$(\mathfrak{A}-\lambda\mathfrak{E})(\mathfrak{T}\mathfrak{v})=\big(\mathfrak{T}(\mathfrak{A}-\lambda\mathfrak{E})\mathfrak{T}^{-1}\big)(\mathfrak{T}\mathfrak{v})=\mathfrak{T}\big((\mathfrak{A}-\lambda\mathfrak{E})\mathfrak{v}\big)=0,$$

d.h. $\mathfrak{T}\mathfrak{v}\in V_\lambda$. Damit haben wir in V_λ einen gegenüber $\mathfrak{G}$ invarianten Untervektorraum von V gefunden und können, da dies für jeden Eigenwert von $\mathfrak{A}$ gilt, V als direkte Summe aller (paarweise orthogonalen und gegen $\mathfrak{G}$ invarianten) Eigenräume von $\mathfrak{A}$ darstellen. Weil nun weiter jeder Eigenraum von $\mathfrak{A}$ wieder direkte Summe von paarweise orthogonalen, bezüglich $\mathfrak{G}$ invarianten und irreduziblen Untervektorräumen ist, erhalten wir so durch Aufspaltung aller Eigenräume eine Summendarstellung von V mit paarweise orthogonalen, irreduziblen, invarianten Untervektorräumen als Summanden. Letztere müssen also wegen der vorausgesetzten Eindeutigkeit einer derartigen Darstellung mit den gegebenen Vektorräumen $W^{(\nu)}$ $(\nu=1,\ldots,s)$ übereinstimmen; d.h. jeder Eigenraum von $\mathfrak{A}$ ist direkte Summe von einigen $W^{(\nu)}$. Wählt man jetzt für V die im Hilfssatz angegebene orthonormierte Basis und nennt die Komponenten eines beliebigen Vektors aus V bezüglich dieser Basis y_t, so nimmt die betrachtete quadratische Form die Gestalt $\mathfrak{y}^*\mathfrak{B}\mathfrak{y}$ an ($\mathfrak{y}$ = Spaltenmatrix mit den Elementen y_t), wobei $\mathfrak{B}$ eine Diagonalmatrix

mit den Diagonalelementen b_t ist $(t=1, \ldots, n_s)$. Hierbei sind alle b_t mit $n_{r-1}<t\leqq n_r$ und $W^{(r)}\subseteq V_\lambda$ gleich dem Eigenwert λ des zu λ gehörenden Eigenraums V_λ von $\mathfrak{A}$, so daß $\mathfrak{y}^*\mathfrak{B}\mathfrak{y}$ in der Tat die in (8) angegebene Gestalt besitzt, womit Hilfssatz 1 bewiesen ist. q.e.d.

Auf Grund von Hilfssatz 1 können nun die Beweise von Satz 1, Satz 2 und Satz 3 auf den Beweis der folgenden Sätze zurückgeführt werden:

SATZ 1'. *Falls nicht gerade $l=m=2$ gilt, ist der Tangentialraum T von $G_{l,m}$ bezüglich der Gruppe $(O(l)\otimes O(m))^*$ irreduzibel. Im Falle $l=m=2$ läßt sich dagegen T auf die einzig mögliche Weise als direkte Summe der beiden orthogonalen und bezüglich $(O(2)\otimes O(2))^*$ invarianten und irreduziblen Untervektorräumen $T^{(1)}$ und $T^{(2)}$ darstellen, welche durch $x_{13}-x_{24}=0$, $x_{23}+x_{14}=0$ und $x_{13}+x_{24}=0$, $x_{23}-x_{14}=0$ gegeben sind.*

SATZ 2'. *T ist bezüglich der Gruppe $O(l)\otimes O(m)$ immer irreduzibel.*

SATZ 3'. *Die Sätze 1' und 2' bleiben wörtlich richtig, wenn darin $(O(l)\otimes O(m))^*$ bzw. $O(l)\otimes O(m)$ durch $SO(l)\otimes SO(m)$ bzw. $SO(l)\otimes O(m)$ ersetzt wird.*

Um dies einzusehen, müssen wir nur noch benutzen, daß im Falle $l=m=2$ alle gegen $(O(2)\otimes O(2))^*$ bzw. $SO(2)\otimes SO(2)$ invarianten, positiv definiten, quadratischen Formen auf T nach Einführung der in Hilfssatz 1 angegebenen Komponententransformation

$$\begin{aligned}\sqrt{2}\,y_1&=x_{13}+x_{24}\\ \sqrt{2}\,y_2&=x_{23}-x_{14}\\ \sqrt{2}\,y_3&=x_{13}-x_{24}\\ \sqrt{2}\,y_4&=x_{23}+x_{14}\end{aligned}$$

die in (8) angegebene Form:

$$\lambda_1\left((y_1)^2+(y_2)^2\right)+\lambda_2\left((y_3)^2+(y_4)^2\right)\qquad(\lambda_1>0,\ \lambda_2>0)$$

annehmen, woraus sich durch Einsetzen ihre ursprüngliche Gestalt:

$$\lambda\left((x_{13})^2+(x_{23})^2+(x_{14})^2+(x_{24})^2\right)+2\mu\,(x_{13}\,x_{24}-x_{23}\,x_{14})$$

mit $\lambda=\frac{1}{2}(\lambda_1+\lambda_2)$ und $\mu=\frac{1}{2}(\lambda_1-\lambda_2)$, d.h. $\lambda_1=\lambda+\mu>0$ und $\lambda_2=\lambda-\mu>0$ oder $\lambda>|\mu|$ ergibt. Diesen quadratischen Formen entsprechen gerade die gegen $SO(4)$ invarianten Riemannschen Metriken (7) von $G_{2,2}$ bzw. $\widetilde{G}_{2,2}$.

Den Beweisen von Satz 1', Satz 2' und Satz 3' schicken wir zunächst weitere Hilfssätze voraus:

HILFSSATZ 2. *Gegeben sei ein endlichdimensionaler Vektorraum V und eine Gruppe $\mathfrak{G}$ von linearen Transformationen von V. Läßt sich dann V sowohl als direkte Summe der bezüglich $\mathfrak{G}$ invarianten und irreduziblen Untervektorräume $V^{(r)}$ $(r=1, \ldots, s)$ als auch als direkte Summen der bezüglich $\mathfrak{G}$ invarianten und irreduziblen Untervektorräume $W^{(r')}$ $(r'=1, \ldots, s')$ darstellen, so gilt $s'=s$, und es existiert eine solche Permutation π der Zahlen $1, \ldots, s$, daß $V^{(r)}$ und $W^{(\pi(r))}$ dieselbe Dimension besitzen $(r=1, \ldots, s)$.*

Einen Beweis dieses Satzes für den Fall eines komplexen Vektorraums V findet man in [4], S. 48; er läßt sich genau so für einen reellen Vektorraum V führen.

HILFSSATZ 3. *Gestattet V die beiden Darstellungen $V=V^{(1)}\dotplus V^{(2)}$ und $V=W^{(1)}\dotplus W^{(2)}$ als direkte Summe der gegen $\mathfrak{G}$ invarianten, irreduziblen Untervektorräume $V^{(1)}$, $V^{(2)}$ und $W^{(1)}$, $W^{(2)}$ mit*

$$\operatorname{Dim} V^{(1)}=\operatorname{Dim} W^{(1)}>\operatorname{Dim} V^{(2)}=\operatorname{Dim} W^{(2)}$$

(vgl. Hilfssatz 2!), *so gilt $V^{(1)}=W^{(1)}$ und $V^{(2)}=W^{(2)}$*[10]).

BEWEIS. Wegen $\operatorname{Dim}(V^{(1)}\cap W^{(1)})\geqq\operatorname{Dim} V^{(1)}+\operatorname{Dim} W^{(1)}-\operatorname{Dim} V=\operatorname{Dim} V^{(1)}-\operatorname{Dim} V^{(2)}>0$ ist $V^{(1)}\cap W^{(1)}$ ein vom Nullraum verschiedener, gegen $\mathfrak{G}$ invarianter Untervektorraum von $V^{(1)}$. Da aber $V^{(1)}$ als bezüglich $\mathfrak{G}$ irreduzibel vorausgesetzt war, gilt $V^{(1)}\cap W^{(1)}=V^{(1)}$, woraus zusammen mit $\operatorname{Dim} V^{(1)}=\operatorname{Dim} W^{(1)}$ $V^{(1)}=W^{(1)}$ folgt. Wir betrachten nun den Verbindungsraum U von $V^{(2)}$ und $W^{(2)}$. U ist gegen alle Transformationen aus $\mathfrak{G}$ invariant, und $\operatorname{Dim} U\leqq\operatorname{Dim} V^{(2)}+\operatorname{Dim} W^{(2)}<\operatorname{Dim} V$. Aus diesem Grunde kann U nicht $V^{(1)}$ enthalten, da U sonst wegen $U\geqq V^{(2)}$ mit V identisch sein müßte. $U\cap V^{(1)}$ ist also als gegen $\mathfrak{G}$ invarianter, echter Untervektorraum des irreduziblen Vektorraums $V^{(1)}$ gleich dem Nullraum; d.h. $0=\operatorname{Dim}(U\cap V^{(1)})\geqq\operatorname{Dim} U+\operatorname{Dim} V^{(1)}-\operatorname{Dim} V$, oder $\operatorname{Dim} U\leqq\operatorname{Dim} V^{(2)}=\operatorname{Dim} W^{(2)}$. Daraus folgt aber zusammen mit $U\geqq V^{(2)}$ und $U\geqq W^{(2)}$: $U=V^{(2)}=W^{(2)}$, und die Eindeutigkeit der in Hilfssatz 3 angegebenen Summendarstellungen ist gezeigt. q.e.d.

HILFSSATZ 4. *Ist $\mathfrak{C}$ eine quadratische l-reihige und $\mathfrak{D}$ eine quadratische m-reihige Matrix, welche beide direkte Summen der quadratischen Matrizen $\mathfrak{C}_1$ und $\mathfrak{C}_2$ bzw. $\mathfrak{D}_1$ und $\mathfrak{D}_2$ sind: $\mathfrak{C}=\mathfrak{C}_1\dotplus\mathfrak{C}_2$, $\mathfrak{D}=\mathfrak{D}_1\dotplus\mathfrak{D}_2$*[11]), *so gilt für das nach* (4) *gebildete Tensorprodukt von $\mathfrak{C}$ und $\mathfrak{D}$, wenn man für die in* (4) *auftretenden Indexkombinationen (i,α) und (j,β) $(i,j=1,\ldots,l;\ \alpha,\beta=l+1,\ldots,n)$ die Reihenfolge $(1,l+1),\ldots,(l,l+1),\ (1,l+2),\ldots,(l,l+2),\ldots,(1,n),\ldots,(l,n)$ festlegt $(n=l+m)$:*

$$\mathfrak{C}\otimes_1(\mathfrak{D}_1\dotplus\mathfrak{D}_2)=(\mathfrak{C}\otimes_1\mathfrak{D}_1)\dotplus(\mathfrak{C}\otimes_1\mathfrak{D}_2). \tag{9}$$

Bei Änderung der angegebenen Reihenfolge in die Reihenfolge $(1,l+1),\ldots,(1,n),$ $(2,l+1),\ldots,(2,n),\ldots,(l,l+1),\ldots,(l,n)$ ist dagegen:

$$(\mathfrak{C}_1\dotplus\mathfrak{C}_2)\otimes_2\mathfrak{D}=(\mathfrak{C}_1\otimes_2\mathfrak{D})\dotplus(\mathfrak{C}_2\otimes_2\mathfrak{D})\text{[12])}. \tag{10}$$

Der Beweis dieses Hilfssatzes erfolgt durch einfache rechnerische Verifikation[13]).

[10]) Es würde für die Gültigkeit dieses Hilfssatzes genügen, nur $V^{(1)}$ als irreduzibel vorauszusetzen.

[11]) Zu dieser Schreibweise vgl. [4], S. 16.

[12]) Mit dem an das Zeichen $\otimes$ angehängten Index wird angedeutet, daß das Tensorprodukt von Matrizen nicht eindeutig definierbar ist, sondern von der Anordnung gewisser Indexkombinationen abhängt.

[13]) Vgl. auch [4], S. 179.

Beweis von Satz 1'.

a) Fall $l=m=2$.

Zunächst bestimmt sich nach (4) für ein beliebiges Element von $(O(2)\otimes O(2))^*$ die Transformationsmatrix bezüglich des Komponentensystems $(x_{13}, x_{23}, x_{14}, x_{24})$ von T zu:

$$(11)\quad \left\{\begin{array}{c} \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\varepsilon\sin\varphi & \varepsilon\cos\varphi \end{pmatrix} \otimes_1 \begin{pmatrix} \cos\psi & \sin\psi \\ -\varepsilon\sin\psi & \varepsilon\cos\psi \end{pmatrix} \\ = \begin{pmatrix} \cos\varphi\cos\psi & \sin\varphi\cos\psi & \cos\varphi\sin\psi & \sin\varphi\sin\psi \\ -\varepsilon\sin\varphi\cos\psi & \varepsilon\cos\varphi\cos\psi & -\varepsilon\sin\varphi\sin\psi & \varepsilon\cos\varphi\sin\psi \\ -\varepsilon\cos\varphi\sin\psi & -\varepsilon\sin\varphi\sin\psi & \varepsilon\cos\varphi\cos\psi & \varepsilon\sin\varphi\cos\psi \\ \sin\varphi\sin\psi & -\cos\varphi\sin\psi & -\sin\varphi\cos\psi & \cos\varphi\cos\psi \end{pmatrix}^{14}). \\ (0\leqq\varphi<2\pi,\ 0\leqq\psi<2\pi,\ \varepsilon=+1 \text{ oder } -1). \end{array}\right.$$

Daraus errechnet sich unmittelbar, daß der Vektor $(a, b, -b, a)$ (a, b = beliebig) aus T bei der zur Matrix (11) gehörenden Transformation in den Vektor

$$(12)\quad \left\{\begin{array}{l} (a\cos(\varphi-\psi)+b\sin(\varphi-\psi),\ -\varepsilon a\sin(\varphi-\psi)+\varepsilon b\cos(\varphi-\psi), \\ \varepsilon a\sin(\varphi-\psi)-\varepsilon b\cos(\varphi-\psi),\ a\cos(\varphi-\psi)+b\sin(\varphi-\psi)) \end{array}\right.$$

übergeht. Dies bedeutet aber, daß der durch die Gleichungen $x_{13}-x_{24}=0$, $x_{23}+x_{14}=0$ gegebene, zweidimensionale Untervektorraum $T^{(1)}$ von T gegenüber $(O(2)\otimes O(2))^*$ invariant ist. Wegen (5) gilt dasselbe auch für das orthogonale Komplement $T^{(2)}$ von $T^{(1)}$ bezüglich T, welches etwa durch die Gleichungen $x_{13}+x_{24}=0$, $x_{23}-x_{14}=0$ bestimmt ist und ebenfalls einen zweidimensionalen Untervektorraum von T darstellt.

Wir zeigen nun, daß $T^{(1)}$ und $T^{(2)}$ bezüglich der in $(O(2)\otimes O(2))^*$ enthaltenen Untergruppe $SO(2)\otimes SO(2)$ irreduzibel sind; dasselbe gilt dann um so mehr auch bezüglich $(O(2)\otimes O(2))^*$ selbst (vgl. Definition 2!). Dies geschehe indirekt; man nimmt zunächst an, $T^{(1)}$ sei bezüglich $SO(2)\otimes SO(2)$ reduzibel. Da $SO(2)\otimes SO(2)$ wegen (5) eine orthogonale Gruppe ist, kann jetzt $T^{(1)}$ als direkte Summe von zwei orthogonalen, eindimensionalen, bezüglich $SO(2)\otimes SO(2)$ invarianten und irreduziblen Untervektorräumen $R^{(1)}$ und $S^{(1)}$ dargestellt werden:

$$(13)\quad T^{(1)}=R^{(1)}\dotplus S^{(1)}.$$

Ist nun $\mathfrak{r}^{(1)}\neq 0$ ein Vektor aus $R^{(1)}$, so gilt für jedes beliebige Element $A\in SO(2)\otimes SO(2)$: $A(\mathfrak{r}^{(1)})=c(A)\mathfrak{r}^{(1)}$, wobei $c(A)$ eine von A stetig abhängende reelle Zahl ist, die wegen $A\in O(4)$ nur die Werte $+1$ und -1 annehmen kann. Da aber $SO(2)\otimes SO(2)$ (bezüglich der natürlichen Topologie) zusammenhängend ist und die identische Transformation enthält, muß $c(A)\equiv+1$

[14]) Es ist bemerkenswert, daß ein beliebiges Element von $SO(2)\otimes SO(2)$ sich auch durch die Quaternionengleichung $x'=(\cos\varphi e-\sin\varphi i)\,x\,(\cos\psi e-\sin\psi j)$ mit $x=x_{13}e+x_{23}i+x_{14}j+x_{24}k$ und $x'=x'_{13}e+x'_{23}i+x'_{14}j+x'_{24}k$ darstellen läßt.

gelten, und wir haben für alle $A \in SO(2) \otimes SO(2)$ die Beziehung:

$$A(\mathfrak{r}^{(1)}) = \mathfrak{r}^{(1)}. \tag{14}$$

Ebenso sieht man ein, daß für alle $A \in SO(2) \otimes SO(2)$ und für einen Vektor $\mathfrak{f}^{(1)} \neq 0$ aus $S^{(1)}$ die Beziehung:

$$A(\mathfrak{f}^{(1)}) = \mathfrak{f}^{(1)} \tag{15}$$

bestehen muß. (14) und (15) ergibt wegen (13) zusammen die Invarianz jedes Vektors von $T^{(1)}$ gegenüber allen Transformationen aus $SO(2) \otimes SO(2)$. Daraus würde insbesondere folgen, daß der Vektor $(1, 0, 0, 1) \in T^{(1)}$ bei $SO(2) \otimes SO(2)$ stets invariant bleibt, während er nach (12) (mit $\varepsilon = +1$!) in den für $\varphi \neq \psi$ von $(1, 0, 0, 1)$ verschiedenen Vektor $(\cos(\varphi - \psi), -\sin(\varphi - \psi), \sin(\varphi - \psi), \cos(\varphi - \psi))$ übergeführt wird. Dieser Widerspruch zeigt, daß $T^{(1)}$ in der Tat bezüglich $SO(2) \otimes SO(2)$ irreduzibel ist. Auf analoge Weise findet man, daß $T^{(2)}$ entweder auch bezüglich $SO(2) \otimes SO(2)$ irreduzibel ist oder bei allen Transformationen von $SO(2) \otimes SO(2)$ elementweise festbleibt. Der letztere Fall kann wiederum nicht eintreten, da der Vektor $(1, 0, 0, -1) \in T^{(2)}$ bei der Transformation mit der Matrix (11) und $\varepsilon = +1$ in den für $\varphi + \psi \neq 0$ und $\varphi + \psi \neq 2\pi$ davon verschiedenen Vektor $(\cos(\varphi + \psi), -\sin(\varphi + \psi), -\sin(\varphi + \psi), -\cos(\varphi + \psi))$ übergeht.

Damit haben wir T als direkte Summe der orthogonalen und bezüglich $(O(2) \otimes O(2))^*$ (wie auch bezüglich $SO(2) \otimes SO(2)$) invarianten und irreduziblen Untervektorräume $T^{(1)}$ und $T^{(2)}$ dargestellt:

$$T = T^{(1)} \dotplus T^{(2)}, \tag{16}$$

und es bleibt nun noch die Satz 1′ behauptete Eindeutigkeit einer derartigen Darstellung nachzuweisen. Hierzu kann man wegen $\mathrm{Dim}\, T^{(1)} = \mathrm{Dim}\, T^{(2)}$ leider nicht den Hilfssatz 3 anwenden, sondern muß anders vorgehen. Dies geschieht, indem man zunächst untersucht, wie sich die Vektorräume $T^{(1)}$ und $T^{(2)}$ verhalten, wenn man sich auf diejenige Untergruppe $[SO(2)]^2$ von $SO(2) \otimes SO(2)$ beschränkt, welche aus den „Tensorquadraten" aller Transformationen aus $SO(2)$ besteht und deshalb durch die Matrizen (11) mit $\varepsilon = +1$ und $\varphi = \psi$ gegeben ist. Wir ersehen jetzt aus (12), daß jeder Vektor von T der Form $(a, b, -b, a)$, d.h. jeder Vektor aus $T^{(1)}$ bei allen Transformationen aus $[SO(2)]^2$ invariant bleibt. Dies bedeutet, anders ausgedrückt:

$$T^{(1)} = U^{(1)} \dotplus U^{(2)}, \tag{17}$$

wobei $U^{(1)}$ und $U^{(2)}$ zwei eindimensionale, bezüglich $[SO(2)]^2$ invariante (und irreduzible) Untervektorräume von $T^{(1)}$ sind. $T^{(2)}$ bleibt dagegen bezüglich $[SO(2)]^2$ irreduzibel, denn sonst müßte wegen des Zusammenhängendseins von $[SO(2)]^2$, wie die schon für $SO(2) \otimes SO(2)$ durchgeführte Überlegung zeigt, jeder Vektor aus $T^{(2)}$ gegenüber allen Transformationen aus $[SO(2)]^2$ invariant bleiben. Dies würde aber wegen der elementweisen Invarianz von $T^{(1)}$ gegen $[SO(2)]^2$ sogar die elementweise Invarianz von ganz T gegen

$[SO(2)]^2$ nach sich ziehen, offensichtlich eine Unmöglichkeit. Aus (16) und (17) resultiert also die Darstellung:

$$T = U^{(1)} \dotplus U^{(2)} \dotplus T^{(2)} \tag{18}$$

von T als direkte Summe von bezüglich $[SO(2)]^2$ invarianten, irreduziblen Untervektorräumen; und $T^{(1)}$ ist dabei genau der Untervektorraum aller derjenigen Vektoren von T, die sich bei allen Transformationen von $[SO(2)]^2$ nicht ändern.

Existierte nun außer (16) noch eine andere „Zerlegung" von T in eine direkte Summe von orthogonalen, bezüglich $(O(2)\otimes O(2))^*$ invarianten, irreduziblen Untervektorräumen, so können hierbei auf Grund von Hilfssatz 2 nur zwei zweidimensionale Summanden auftreten: $T = T'^{(1)} \dotplus T'^{(2)}$. Beschränkt man sich daraufhin auf $[SO(2)]^2$, so muß sich nach (18) und Hilfssatz 2 einer dieser beiden Summanden, etwa $T'^{(1)}$, weiter in eine direkte Summe zweier eindimensionaler, gegenüber $[SO(2)]^2$ invarianter Untervektorräume zerlegen lassen. Dann folgt aus dem Zusammenhängendsein von $[SO(2)]^2$ wieder die elementweise Invarianz von $T'^{(1)}$ gegenüber $[SO(2)]^2$, d.h. wegen der vorhin erwähnten Eigenschaft von $T^{(1)}$ in Verbindung mit $\mathrm{Dim}\, T'^{(1)} = \mathrm{Dim}\, T^{(1)}$ gilt: $T'^{(1)} = T^{(1)}$. Weil $T'^{(2)}$ bzw. $T^{(2)}$ die orthogonalen Komplemente von $T'^{(1)}$ bzw. $T^{(1)}$ in bezug auf T sind, ergibt sich aus $T'^{(1)} = T^{(1)}$ schließlich auch $T'^{(2)} = T^{(2)}$; und mit der gezeigten Eindeutigkeit der Darstellung (16) ist Satz 1′ für den Fall $l = m = 2$ vollständig bewiesen.

Wir bemerken noch, daß man ganz analog auch die Eindeutigkeit der Darstellung (16) *bezüglich der Gruppe* $SO(2)\otimes SO(2)$ *einsieht.* Dies zieht eine wichtige Konsequenz nach sich; wir behaupten: *T ist sowohl bezüglich* $SO(2)\otimes O(2)$ *als auch bezüglich* $O(2)\otimes SO(2)$ *irreduzibel.* Wäre nämlich etwa T bezüglich $SO(2)\otimes O(2)$ reduzibel, so müßte T wegen $SO(2)\otimes O(2) \supsetneqq SO(2)\otimes SO(2)$ und dem Ebengesagten direkte Summe der gegen $SO(2)\otimes O(2)$ invarianten Untervektorräume $T^{(1)}$ und $T^{(2)}$ sein. Die Transformation aus $SO(2)\otimes O(2)$ mit der Matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes_1 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

führt jedoch den Vektor $(1, 0, 0, 1) \in T^{(1)}$ in den Vektor $(0, 1, 1, 0) \notin T^{(1)}$ über, und dieser Widerspruch beweist die Irreduzibilität von T bezüglich $SO(2)\otimes O(2)$. Analog zeigt man auch die Irreduzibilität von T bezüglich $O(2)\otimes SO(2)$.

b) Fall $l \neq 2$ oder $m \neq 2$.

Wir beweisen nun die Irreduzibilität von T bezüglich $SO(l)\otimes SO(m)$, falls nicht $l = m = 2$; dann gilt dasselbe um so mehr auch bezüglich der Gruppe $(O(l)\otimes O(m))^*$, wie in Satz 1′ behauptet wurde. Dieser Nachweis geschieht durch vollständige Induktion nach $l + m = n$. Für $2 \leqq n \leqq 4$ ist die Behauptung trivial, da dann $l = 1$ oder $m = 1$ gilt ($l = m = 2$ war ja ausgeschlossen!) und

somit $SO(l)\otimes SO(m)$ mit den Gruppen $SO(m)$ oder $SO(l)$ isomorph ist, bezüglich derer T gewiß irreduzibel ist. Wir nehmen jetzt also an, unsere Behauptung sei für $n-1$ ($n\geqq 5$) bewiesen, und werden daraus ihre Gültigkeit für n folgern.

Dazu betrachten wir zunächst den Fall $l\leqq m$ (d.h. $m\geqq 3$). $SO(l)\otimes SO(m)$ enthält die Gruppe $SO(l)\otimes(O(1)\dotplus O(m-1))^*$, wobei $(O(1)\dotplus O(m-1))^*$ diejenige Untergruppe von $SO(m)$ bedeuten soll, deren Elemente durch Matrizen der Form $(\varepsilon)\dotplus\mathfrak{D}$[15]) ($\mathfrak{D}$ = beliebige $(m-1)$-reihige orthogonale Matrix, ε = Wert der Determinante von $\mathfrak{D}$) gegeben sind. Auf Grund von (9) stellt man leicht fest, daß der lm-dimensionale Tangentialraum T von $G_{l,m}$ direkte Summe der gegen $SO(l)\otimes((O(1)\dotplus O(m-1))^*$ invarianten l-dimensionalen und $l(m-1)$-dimensionalen Untervektorräume $R^{(1)}$: $x_{i\alpha'}=0$ ($i=1,\ldots,l$; $\alpha'=l+2,\ldots,n$) und $S^{(1)}$: $x_{il+1}=0$ ($i=1;\ldots,l$) ist. $R^{(1)}$ und $S^{(1)}$ sind sogar bezüglich $SO(l)\otimes(O(1)\dotplus O(m-1))^*$ irreduzibel, weil $R^{(1)}$ bezüglich $SO(l)\otimes O(1)$ und $S^{(1)}$ bezüglich $SO(l)\otimes O(m-1)$ irreduzibel ist. Das erstere ist trivial und das letztere folgt wegen $SO(l)\otimes O(m-1)\supseteq SO(l)\otimes SO(m-1)$ aus der Induktionsvoraussetzung bzw. im Fall $l=2$, $m=3$ aus der auf S. 345 festgestellten Irreduzibilität von $S^{(1)}$ bezüglich $SO(2)\otimes O(2)$. Wäre nun T bezüglich $SO(l)\otimes SO(m)$ reduzibel, so auch bezüglich $SO(l)\otimes(O(1)\dotplus O(m-1))^*$, und T müßte auf Grund von Hilfssatz 2 direkte Summe eines l-dimensionalen und eines $l(m-1)$-dimensionalen, sowohl bezüglich $SO(l)\otimes SO(m)$ als auch bezüglich $SO(l)\otimes(O(1)\dotplus O(m-1))^*$ invarianten und irreduziblen Untervektorraums sein. Wegen $m>2$, d.h. $l(m-1)>l$ stimmten dann nach Hilfssatz 3 ($\mathfrak{G}=SO(l)\otimes(O(1)\dotplus O(m-1))^*$!) diese beiden Vektorräume mit den angegebenen Räumen $R^{(1)}$ und $S^{(1)}$ überein, so daß die einzig mögliche Zerlegung von T in bezüglich $SO(l)\otimes SO(m)$ invariante, irreduzible Untervektorräume lautet:

$$T=R^{(1)}\dotplus S^{(1)}. \tag{19}$$

Andererseits enthält $SO(l)\otimes SO(m)$ aber auch die Gruppe $SO(l)\otimes(O(m-1)\dotplus O(1))^*$, wobei jetzt $(O(m-1)\dotplus O(1))^*$ die Gruppe aller Transformationen mit Matrizen der Form $\mathfrak{D}\dotplus(\varepsilon)$ ($\mathfrak{D}=(m-1)$-reihige orthogonale Matrix, ε = Wert der Determinante von $\mathfrak{D}$) darstellt. Man findet genau wie im letzten Abschnitt unter Benutzung von Hilfssatz 2, Hilfssatz 3 und (9)

$$T=R^{(2)}\dotplus S^{(2)} \tag{20}$$

als die einzig mögliche Zerlegung von T in bezüglich $SO(l)\otimes SO(m)$ invariante, irreduzible Untervektorräume; hierbei ist aber $R^{(2)}$ durch $x_{i\alpha''}=0$ ($i=1,\ldots,l$; $\alpha''=l+1,\ldots,n-1$) und $S^{(2)}$ durch $x_{in}=0$ ($i=1,\ldots,l$) definiert. Nun sind die Zerlegungen (19) und (20) wegen $m>2$, d.h. $n>l+1$ gewiß verschieden; und somit kann T bezüglich $SO(l)\otimes SO(m)$ ($l\leqq m$) nicht reduzibel sein. Dasselbe trifft auch im Falle $l\geqq m$ zu, wie man durch Einschränkung von $SO(l)\otimes SO(m)$ auf die (entsprechend definierten) Untergruppen

[15]) Vgl. Fußnote 11.

$(O(1)\dotplus O(l-1))^*\otimes SO(m)$ bzw. $(O(l-1)\dotplus O(1))^*\otimes SO(m)$ unter Benutzung der Hilfssätze 2, 3 und 4 (10) sowie der Irreduzibilität des Tangentialraums bezüglich $O(2)\otimes SO(2)$ einsieht. Also gilt die Induktionsbehauptung (Irreduzibilität von T bezüglich $SO(l)\otimes SO(m)$, wenn nicht $l=m=2$) in jedem Fall, und unser Induktionsbeweis wie der Beweis von Satz 1′ ist zu Ende geführt. q.e.d.

Beweise von Satz 2′ und Satz 3′.

Der Tangentialraum T ist bezüglich $O(l)\otimes O(m)$ stets irreduzibel; dies folgt im Spezialfall $l=m=2$ aus der beim Beweis von Satz 1′ festgestellten Irreduzibilität von T bezüglich $SO(2)\otimes O(2)$ und im allgemeinen Falle aus der ebenfalls gezeigten Irreduzibilität von T bezüglich $SO(l)\otimes SO(m)$. Damit ist Satz 2′ schon bewiesen, und wir wenden uns dem Beweis von Satz 3′ zu. Dieser Satz sagt im einzelnen aus: Für $l=m=2$ besitzt T bezüglich $SO(2)\otimes SO(2)$ die einzig mögliche Zerlegung $T=T^{(1)}\dotplus T^{(2)}$[16]), sonst ist T bezüglich $SO(l)\otimes SO(m)$ irreduzibel, während T bezüglich $SO(l)\otimes O(m)$ für beliebige l und m irreduzibel ist. Die ersten beiden Aussagen hiervon wurden beim Beweis von Satz 1′ schon gezeigt, und die letzte Aussage ergibt sich wiederum aus der bewiesenen Irreduzibilität von T bezüglich $SO(2)\otimes O(2)$ bzw. bezüglich $SO(l)\otimes SO(m)$ $(l\neq 2$ oder $m\neq 2)$.

Hiermit sind die Sätze 1′, 2′ und 3′ und somit auch die Sätze 1, 2 und 3 vollständig bewiesen; wir bemerken noch, daß in unseren Beweisen von der Theorie der Lieschen Gruppen kein Gebrauch gemacht wurde.

3. Krümmungseigenschaften der mit invarianten Riemannschen Metriken versehenen Grassmannschen Mannigfaltigkeiten

3.1. Zunächst sollen in diesem Abschnitt der Riemannsche Krümmungstensor der Metrik (6) von $G_{l,m}$ explizit berechnet und daraus einige Folgerungen für die Krümmungseigenschaften von $G_{l,m}$ gezogen werden. Wegen der Invarianz der betrachteten Riemannschen Metrik gegenüber der auf $G_{l,m}$ transitiv wirkenden Gruppe $O(n)$ genügt es hierbei, sich auf einen speziellen Punkt $o\in G_{l,m}$ zu beschränken. Wir nehmen nun an, $u_{i\alpha}$ $(i=1,\ldots,l;\ \alpha=l+1,\ldots,n)$ seien die durch das orthonormierte n-Bein $\{\mathring{e}_j,\ \mathring{e}_\beta\}$ bestimmten, lokalen Koordinaten von $G_{l,m}$ in einer Umgebung $\mathfrak{U}$ von $o=(0,\ldots,0)$ und jeder Punkt $\mathfrak{p}$ dieser Umgebung mit den Koordinaten $u_{i\alpha}$ werde durch das orthonormierte l-Bein $\{\mathfrak{e}_j(u_{i\alpha})\}$ $(j=1,\ldots,l;\ \beta=l+1,\ldots,n)$ dargestellt (vgl. Abschnitt 1!). Da man sich die Vektoren $\mathfrak{e}_j$ durch Orthogonalisierung aus den in (1) definierten Vektoren $\mathfrak{f}_i$ entstanden denken kann, darf ohne Beschränkung der Allgemeinheit angenommen werden, die Vektorfunktionen $\mathfrak{e}_j=\mathfrak{e}_j(u_{i\alpha})$ seien beliebig oft stetig differenzierbar, es soll außerdem $\mathfrak{e}_j(0,\ldots,0)=\mathring{e}_j$ gelten. Ebenfalls kann man annehmen, das l-Bein $\{\mathfrak{e}_j(u_{i\alpha})\}$ werde durch das orthonormierte m-Bein $\{\mathfrak{e}_\beta(u_{i\alpha})\}$ $(\beta=l+1,\ldots,n)$ zu einem orthonormierten n-Bein ergänzt, dessen Vektoren $\mathfrak{e}_\beta(u_{i\alpha})$ auch beliebig oft stetig differenzierbar von den Koordinaten $u_{i\alpha}$ abhängen und für die $\mathfrak{e}_\beta(0,\ldots,0)=\mathring{e}_\beta$ ist. Dann lautet das Linienelement-

[16]) $T^{(1)}$ und $T^{(2)}$ haben hier dieselbe Bedeutung wie in (16).

quadrat der Metrik (6) von $G_{l,n}$:

$$ds^2 = \sum_{i,j=1}^{l} \sum_{\alpha,\beta=l+1}^{n} g_{(i\alpha)(j\beta)}(u_{k\gamma})\, du_{i\alpha}\, du_{j\beta} \tag{21}$$

mit

$$g_{(i\alpha)(j\beta)} = \sum_{k=1}^{l} \sum_{\gamma=l+1}^{n} \left\langle e_\gamma \frac{\partial e_k}{\partial u_{i\alpha}} \right\rangle \left\langle e_\gamma \frac{\partial e_k}{\partial u_{j\beta}} \right\rangle^{17}). \tag{22}$$

Zur Berechnung des Krümmungstensors dieser Metrik erweist sich nun der Cartansche Differentialformenkalkül als besonders geeignet. Definiert man nämlich auf $\mathfrak{U}$ die Pfaffschen Formen $\omega_{(i\alpha)}$ mit Hilfe der Vektordifferentialformen

$$de_i = \sum_{j=1}^{l} \sum_{\beta=l+1}^{n} \frac{\partial e_i}{\partial u_{j\beta}}\, du_{j\beta}, \qquad de_\alpha = \sum_{j=1}^{l} \sum_{\beta=l+1}^{n} \frac{\partial e_\alpha}{\partial u_{j\beta}}\, du_{j\beta}$$

durch

$$\begin{cases} \omega_{(i\alpha)} = \langle e_\alpha\, de_i \rangle = \sum_{j=1}^{l} \sum_{\beta=l+1}^{n} \left\langle e_\alpha \frac{\partial e_i}{\partial u_{j\beta}} \right\rangle du_{j\beta} = -\langle e_i\, de_\alpha \rangle = -\omega_{(\alpha i)}, \\ (i = 1, \ldots, l;\ \alpha = l+1, \ldots, n), \end{cases} \tag{23}$$

so folgt aus (21) und (22):

$$ds^2 = \sum_{i=1}^{l} \sum_{\alpha=l+1}^{n} (\omega_{(i\alpha)})^2 \quad \text{(vgl. (6)!)}.$$

Nach CHERN[18]) existieren jetzt auf $\mathfrak{U}$ stets Pfaffsche Formen $\omega_{(i\alpha)(j\beta)}$ $(i, j = 1, \ldots, l;\ \alpha, \beta = l+1, \ldots, n)$ mit den sie eindeutig bestimmenden Eigenschaften:

$$\omega_{(i\alpha)(j\beta)} = -\omega_{(j\beta)(i\alpha)} \tag{24}$$

und

$$d\omega_{(i\alpha)} = -\sum_{j=1}^{l} \sum_{\beta=l+1}^{n} \omega_{(i\alpha)(j\beta)} \wedge \omega_{(j\beta)}, \tag{25}$$

und die sich aus der Darstellung

$$d\omega_{(i\alpha)(j\beta)} + \sum_{k=1}^{l} \sum_{\gamma=l+1}^{n} \omega_{(i\alpha)(k\gamma)} \wedge \omega_{(k\gamma)(j\beta)} = \tfrac{1}{2} \sum_{k,h=1}^{l} \sum_{\gamma,\delta=l+1}^{n} Q_{(i\alpha)(j\beta)(k\gamma)(h\delta)} (\omega_{(k\gamma)} \wedge \omega_{(h\delta)}) \tag{26}$$

mit

$$Q_{(i\alpha)(j\beta)(k\gamma)(h\delta)} = -Q_{(i\alpha)(j\beta)(h\delta)(k\gamma)}$$

eindeutig ergebenden Koeffizienten $Q_{(i\alpha)(j\beta)(k\gamma)(h\delta)}$ stimmen im Punkt o wegen

$$\omega_{(i\alpha)}(0, \ldots, 0) = du_{i\alpha} \tag{27}$$

mit den entsprechenden Komponenten $R_{(i\alpha)(j\beta)(k\gamma)(h\delta)}$ des Riemannschen Krümmungstensors unserer Metrik überein $(i, j, k, h = 1, \ldots, l;\ \alpha, \beta, \gamma, \delta = l+1, \ldots, n)$.

Diese Tatsachen lassen in unserem Fall die explizite Berechnung von $R_{(i\alpha)(j\beta)(k\gamma)(h\delta)}(0, \ldots, 0)$ folgendermaßen vonstatten gehen: Die Differentiation von (23) ergibt:

$$d\omega_{(i\alpha)} = \sum_{j,k=1}^{l} \sum_{\beta,\gamma=l+1}^{n} \left\langle \frac{\partial e_\alpha}{\partial u_{k\gamma}} \frac{\partial e_i}{\partial u_{j\beta}} \right\rangle du_{k\gamma} \wedge du_{j\beta} = \langle de_\alpha \wedge de_i \rangle,$$

[17]) Die Metrik (21) hängt nicht von der speziellen Wahl der durch die Punkte p von $G_{l,n}$ noch nicht vollständig bestimmten Vektorfunktionen e_j und e_β ab, da sie gegenüber $O(n)$ invariant ist.

[18]) Siehe [5], S. 34–38.

woraus durch Einsetzen von

$$(28)\qquad d\mathfrak{e}_i = \sum_{j=1}^{l} \langle \mathfrak{e}_j\, d\mathfrak{e}_i\rangle\, \mathfrak{e}_j + \sum_{\beta=l+1}^{n} \langle \mathfrak{e}_\beta\, d\mathfrak{e}_i\rangle\, \mathfrak{e}_\beta \qquad (i = 1, \ldots, l)$$

und

$$(29)\qquad d\mathfrak{e}_\alpha = \sum_{j=1}^{l} \langle \mathfrak{e}_j\, d\mathfrak{e}_\alpha\rangle\, \mathfrak{e}_j + \sum_{\beta=l+1}^{n} \langle \mathfrak{e}_\beta\, d\mathfrak{e}_\alpha\rangle\, \mathfrak{e}_\beta \qquad (\alpha = l+1, \ldots, n)^{19})$$

mit den zu (23) hinzukommenden Abkürzungen

$$(30)\qquad \omega_{(ij)} = \langle \mathfrak{e}_j\, d\mathfrak{e}_i\rangle = -\langle \mathfrak{e}_i\, d\mathfrak{e}_j\rangle = -\omega_{(ji)}$$

und

$$(31)\qquad \omega_{(\alpha\beta)} = \langle \mathfrak{e}_\beta\, d\mathfrak{e}_\alpha\rangle = -\langle \mathfrak{e}_\alpha\, d\mathfrak{e}_\beta\rangle = -\omega_{(\beta\alpha)}:$$

$$d\omega_{(i\alpha)} = \sum_{j=1}^{l} \omega_{(\alpha j)} \wedge \omega_{(ij)} + \sum_{\beta=l+1}^{n} \omega_{(\alpha\beta)} \wedge \omega_{(i\beta)} = \sum_{j=1}^{l} \sum_{\beta=l+1}^{n} \left(\omega_{(ij)}\, \delta_{\alpha\beta} + \delta_{ij}\, \omega_{(\alpha\beta)}\right) \wedge \omega_{(j\beta)}$$

folgt. Durch Vergleich von der letzten Beziehung und von

$$\omega_{(ji)}\, \delta_{\beta\alpha} + \delta_{ji}\, \omega_{(\beta\alpha)} = -\left(\omega_{(ij)}\, \delta_{\alpha\beta} + \delta_{ij}\, \omega_{(\alpha\beta)}\right)$$

mit (25) und (24) schließt man auf:

$$\omega_{(i\alpha)(j\beta)} = -\left(\omega_{(ij)}\, \delta_{\alpha\beta} + \delta_{ij}\, \omega_{(\alpha\beta)}\right).$$

Damit wird nach (30), (31), (28), (29), (23):

$$\begin{aligned}
&d\omega_{(i\alpha)(j\beta)} + \sum_{k=1}^{l} \sum_{\gamma=l+1}^{n} \omega_{(i\alpha)(k\gamma)} \wedge \omega_{(k\gamma)(j\beta)} = \langle d\mathfrak{e}_i \wedge d\mathfrak{e}_j\rangle\, \delta_{\alpha\beta} + \delta_{ij} \langle d\mathfrak{e}_\alpha \wedge d\mathfrak{e}_\beta\rangle + \\
&\quad + \sum_{k=1}^{l} \sum_{\gamma=l+1}^{n} \left(\omega_{(ik)}\, \delta_{\alpha\gamma} + \delta_{ik}\, \omega_{(\alpha\gamma)}\right) \wedge \left(\omega_{(kj)}\, \delta_{\gamma\beta} + \delta_{kj}\, \omega_{(\gamma\beta)}\right) \\
&= \Big(\sum_{k=1}^{l} \omega_{(ik)} \wedge \omega_{(jk)} + \sum_{\gamma=l+1}^{n} \omega_{(i\gamma)} \wedge \omega_{(j\gamma)}\Big)\, \delta_{\alpha\beta} + \delta_{ij} \Big(\sum_{k=1}^{l} \omega_{(\alpha k)} \wedge \omega_{(\beta k)} + \sum_{\gamma=l+1}^{n} \omega_{(\alpha\gamma)} \wedge \omega_{(\beta\gamma)}\Big) + \\
&\quad + \Big(\sum_{k=1}^{l} \omega_{(ik)} \wedge \omega_{(kj)}\Big)\, \delta_{\alpha\beta} + \delta_{ij} \Big(\sum_{\gamma=l+1}^{n} \omega_{(\alpha\gamma)} \wedge \omega_{(\gamma\beta)}\Big) + \omega_{(ij)} \wedge \omega_{(\alpha\beta)} + \omega_{(\alpha\beta)} \wedge \omega_{(ij)} \\
&= \Big(\sum_{\gamma=l+1}^{n} \omega_{(i\gamma)} \wedge \omega_{(j\gamma)}\Big)\, \delta_{\alpha\beta} + \delta_{ij} \Big(\sum_{k=1}^{l} \omega_{(k\alpha)} \wedge \omega_{(k\beta)}\Big) \\
&= \sum_{k,h=1}^{l} \sum_{\gamma,\delta=l+1}^{n} \left(\delta_{ik}\, \delta_{jh}\, \delta_{\gamma\delta}\, \delta_{\alpha\beta} + \delta_{ij}\, \delta_{kh}\, \delta_{\alpha\gamma}\, \delta_{\beta\delta}\right) \left(\omega_{(k\gamma)} \wedge \omega_{(h\delta)}\right) \\
&= \tfrac{1}{2} \sum_{k,h=1}^{l} \sum_{\gamma,\delta=l+1}^{n} \left(\left(\delta_{ik}\, \delta_{jh} - \delta_{ih}\, \delta_{jk}\right) \delta_{\alpha\beta}\, \delta_{\gamma\delta} + \delta_{ij}\, \delta_{kh} \left(\delta_{\alpha\gamma}\, \delta_{\beta\delta} - \delta_{\alpha\delta}\, \delta_{\beta\gamma}\right)\right) \left(\omega_{(k\gamma)} \wedge \omega_{(h\delta)}\right) \\
&\qquad\qquad (1 \leqq i, j \leqq l;\ l+1 \leqq \alpha, \beta \leqq n).
\end{aligned}$$

Vergleicht man dieses Resultat mit der allgemeinen Formel (26) und berücksichtigt man dabei die schiefe Symmetrie des eben berechneten Koeffizienten von $\omega_{(k\gamma)} \wedge \omega_{(h\delta)}$ bezüglich der Indexpaare $(k\,\gamma)$ und $(h\,\delta)$, so erhält man schließlich für die Komponenten des Riemannschen Krümmungstensors der Metrik (21):

$$(32)\quad R_{(i\alpha)(j\beta)(k\gamma)(h\delta)}(0, \ldots, 0) = \left(\delta_{ik}\, \delta_{jh} - \delta_{ih}\, \delta_{jk}\right) \delta_{\alpha\beta}\, \delta_{\gamma\delta} + \delta_{ij}\, \delta_{kh} \left(\delta_{\alpha\gamma}\, \delta_{\beta\delta} - \delta_{\alpha\delta}\, \delta_{\beta\gamma}\right).$$

[19]) Hierbei sind die Differentialformen $\langle \mathfrak{e}_j\, d\mathfrak{e}_i\rangle$ und $\langle \mathfrak{e}_\beta\, d\mathfrak{e}_\alpha\rangle$ analog wie die Differentialformen von (23) definiert.

Daraus berechnet sich jetzt im Fall $l\,m>1$ mit der aus (22), (23) und (27) folgenden Beziehung

$$g_{(i\alpha)(j\beta)}(0,\ldots,0)=\delta_{ij}\,\delta_{\alpha\beta} \tag{33}$$

für die Riemannsche Krümmung $K(\mathfrak{x},\mathfrak{y})$ von (21), welche zu dem von den linear unabhängigen Vektoren $\mathfrak{x}=(x_{i\alpha})$ und $\mathfrak{y}=(y_{i\alpha})$ aus dem Tangentialraum T von $G_{l,m}$ im Nullpunkt aufgespannten Flächenelement gehört[20]):

$$K(\mathfrak{x},\mathfrak{y})=\frac{\sum\limits_{i,j,k,h=1}^{l}\ \sum\limits_{\alpha,\beta,\gamma,\delta=l+1}^{n} R_{(i\alpha)(j\beta)(k\gamma)(h\delta)}(0,\ldots,0)\,x_{i\alpha}\,y_{j\beta}\,x_{k\gamma}\,y_{h\delta}}{\sum\limits_{i,j,k,h=1}^{l}\ \sum\limits_{\alpha,\beta,\gamma,\delta=l+1}^{n}\left(g_{(i\alpha)(k\gamma)}\,g_{(j\beta)(h\delta)}-g_{(i\alpha)(h\delta)}\,g_{(j\beta)(k\gamma)}\right)(0,\ldots,0)\,x_{i\alpha}\,y_{j\beta}\,x_{k\gamma}\,y_{h\delta}}$$

$$=\frac{\sum\limits_{\alpha,\gamma=l+1}^{n}\ \sum\limits_{i,j=1}^{l}\left(x_{i\alpha}\,y_{j\alpha}\,x_{i\gamma}\,y_{j\gamma}-x_{i\alpha}\,y_{j\alpha}\,x_{j\gamma}\,y_{i\gamma}\right)+\sum\limits_{i,k=1}^{l}\ \sum\limits_{\alpha,\beta=l+1}^{n}\left(x_{i\alpha}\,y_{i\beta}\,x_{k\alpha}\,y_{k\beta}-x_{i\alpha}\,y_{i\beta}\,x_{k\beta}\,y_{k\alpha}\right)}{\sum\limits_{i,j=1}^{l}\ \sum\limits_{\alpha,\beta=l+1}^{n}\left((x_{i\alpha})^2\,(y_{j\beta})^2-x_{i\alpha}\,y_{i\alpha}\,x_{j\beta}\,y_{j\beta}\right)}$$

$$=\frac{\sum\limits_{\substack{\alpha,\gamma\\=l+1}}^{n}\left(\sum\limits_{i,j=1}^{}\left(x_{i\alpha}\,y_{j\alpha}-x_{j\alpha}\,y_{i\alpha}\right)\left(x_{i\gamma}\,y_{j\gamma}-x_{j\gamma}\,y_{i\gamma}\right)\right)+\sum\limits_{i,k=1}^{l}\left(\sum\limits_{\substack{\alpha,\beta\\=l+1}}^{n}\left(x_{i\alpha}\,y_{i\beta}-x_{i\beta}\,y_{i\alpha}\right)\left(x_{k\alpha}\,y_{k\beta}-x_{k\beta}\,y_{k\alpha}\right)\right)}{\sum\limits_{i,j=1}^{l}\ \sum\limits_{\alpha,\beta=l+1}^{n}\left((x_{i\alpha})^2\,(y_{j\beta})^2+(x_{j\beta})^2\,(y_{i\alpha})^2-2\,x_{i\alpha}\,y_{j\beta}\,x_{j\beta}\,y_{i\alpha}\right)}$$

$$=\frac{\sum\limits_{i,j=1}^{l}\left(\sum\limits_{\alpha=l+1}^{n}\left(x_{i\alpha}\,y_{j\alpha}-x_{j\alpha}\,y_{i\alpha}\right)\right)^2+\sum\limits_{\alpha,\beta=l+1}^{n}\left(\sum\limits_{i=1}^{l}\left(x_{i\alpha}\,y_{i\beta}-x_{i\beta}\,y_{i\alpha}\right)\right)^2}{\sum\limits_{i,j=1}^{l}\ \sum\limits_{\alpha,\beta=l+1}^{n}\left(x_{i\alpha}\,y_{j\beta}-x_{j\beta}\,y_{i\alpha}\right)^2}.$$

$K(\mathfrak{x},\mathfrak{y})$ ist also stets nichtnegativ; und im Fall $l=1$, $m>1$ oder $m=1$, $l>1$ sogar vom konstanten Wert 1. In den übrigen Fällen $l>1$ und $m>1$ ist die Riemannsche Krümmung dagegen nicht konstant. Definiert man nämlich $\mathfrak{x}^{(1)}$ und $\mathfrak{y}^{(1)}$ durch $\mathfrak{x}^{(1)}=\{x^{(1)}_{i\alpha}\}$ mit $x^{(1)}_{i\alpha}=\delta_{i1}\,\delta_{\alpha l+1}$ und $\mathfrak{y}^{(1)}=\{y^{(1)}_{i\alpha}\}$ mit $y^{(1)}_{i\alpha}=\delta_{i1}\,\delta_{\alpha l+2}$ $(i=1,\ldots,l;\,\alpha=l+1,\ldots,n)$, so wird $K(\mathfrak{x}^{(1)},\mathfrak{y}^{(1)})=1$, während für $\mathfrak{x}^{(2)}=\{\delta_{i1}\,\delta_{\alpha l+1}\}$ und $\mathfrak{y}^{(2)}=\{\delta_{i2}\,\delta_{\alpha l+2}\}$ $K(\mathfrak{x}^{(2)},\mathfrak{y}^{(2)})=0$ gilt. Daraus folgt gleichzeitig, daß die Riemannsche Krümmung der Metrik (21) nur dann überall positiv sein kann, wenn sie konstant ist.

Der Ricci-Tensor $R_{(j\beta)(k\gamma)}=g^{(i\alpha)(h\delta)}\,R_{(i\alpha)(j\beta)(k\gamma)(h\delta)}$ unserer Metrik hat nach (32) und (33) im Nullpunkt die Komponenten:

$$\begin{cases} R_{(j\beta)(k\gamma)}(0,\ldots,0)=\sum\limits_{i=1}^{l}\ \sum\limits_{\alpha=l+1}^{n} R_{(i\alpha)(j\beta)(k\gamma)(i\alpha)}(0,\ldots,0)=(\delta_{jk}-l\,\delta_{jk})\,\delta_{\beta\gamma}+ \\ \qquad +\,\delta_{jk}(\delta_{\beta\gamma}-m\,\delta_{\beta\gamma})=(2-n)\,\delta_{jk}\,\delta_{\beta\gamma}=(2-n)\,g_{(j\beta)(k\gamma)}(0,\ldots,0), \end{cases} \tag{34}$$

er erweist sich also als zum Maßtensor proportional. Die Grassmannsche Mannigfaltigkeit $G_{l,m}$ stellt deshalb bezüglich der Riemannschen Metrik (21)

[20]) Vgl. [6], S. 81.

einen sog. „Einsteinschen Raum" dar[21]) und besitzt wegen (34) die konstante Skalarkrümmung

$$R = g^{(j\beta)(k\gamma)} R_{(j\beta)(k\gamma)} = g^{(j\beta)(k\gamma)} (2-n)\, g_{(j\beta)(k\gamma)} = \big(2-(l+m)\big)\, l\, m.$$

Berücksichtigt man nun noch den konstanten Maßfaktor der nach Satz 2 durch Invarianz gegenüber $O(n)$ ausgezeichneten Riemannschen Metrik von $G_{l,m}$, so kann man die bisherigen Ergebnisse dieses Abschnitts zusammenfassen in

Satz 4. *Die auf der Grassmannschen Mannigfaltigkeit* $G_{l,m}$[22]) *definierte Riemannsche Metrik* $ds^2 = \lambda\Big(\sum_{i=1}^{l}\sum_{\alpha=l+1}^{n}\big(\langle e_\alpha\, de_i\rangle\big)^2\Big)$ $(\lambda = \text{const} > 0)$ *macht* $G_{l,m}$ *zu einem Einsteinschen Raum mit der konstanten Skalarkrümmung* $R = \frac{1}{\lambda}\big((2-(l+m))\, l\, m\big)$ $(l\, m > 1)$. *Die zu dieser „Normalmetrik" gehörende Riemannsche Krümmung K ist nichtnegativ; sie ist aber nur dann überall positiv, wenn sie konstant ist. Dabei besitzt* $G_{l,m}$ *genau dann konstantes K, wenn* $l=1$, $m>1$ *oder* $m=1$, $l>1$ *gilt; in diesen Fällen hat K den Wert* $1/\lambda$.

3.2. Wir wollen jetzt entsprechende Resultate für die nach Satz 1 neben der „Normalmetrik" einzige weitere, ebenfalls gegenüber $SO(n)$ invariante Riemannsche Metrik einer Grassmannschen Mannigfaltigkeit, nämlich für die auf $G_{2,2}$ definierte „Ausnahmemetrik" (7) mit $(\lambda > |\mu| > 0!)$ herleiten. Zu diesem Zweck führen wir die Differentialformen

$$\omega_1 = \sqrt{\frac{\lambda_1}{2}}\,\big(\langle e_3\, de_1\rangle + \langle e_4\, de_2\rangle\big)$$

$$\omega_2 = \sqrt{\frac{\lambda_1}{2}}\,\big(\langle e_3\, de_2\rangle - \langle e_4\, de_1\rangle\big)$$

$$\omega_3 = \sqrt{\frac{\lambda_2}{2}}\,\big(\langle e_3\, de_1\rangle - \langle e_4\, de_2\rangle\big)$$

$$\omega_4 = \sqrt{\frac{\lambda_2}{2}}\,\big(\langle e_3\, de_2\rangle + \langle e_4\, de_1\rangle\big)$$[23])

mit

$$0 < \lambda_1 = \lambda + \mu = \text{const}, \qquad 0 < \lambda_2 = \lambda - \mu = \text{const}. \tag{35}$$

ein, wodurch (7) die Form

$$ds^2 = \sum_{a=1}^{4} (\omega_a)^2$$

annimmt (vgl. S. 341!). Der Cartansche Kalkül ergibt jetzt analog wie bei der Normalmetrik:

$$d\omega_a = -\sum_{b=1}^{4} \omega_{ab} \wedge \omega_b$$

[21]) Siehe [6], S. 92, Übg. 2ff.

[22]) Natürlich gilt Satz 4 in derselben Form auch für die entsprechend auf $\widetilde{G_{l,m}}$ definierte Normalmetrik, da $\widetilde{G_{l,m}}$ eine zweifache reguläre Überlagerung von $G_{l,m}$ darstellt.

[23]) Die Differentialformen $\langle e_\alpha\, de_i\rangle$; $\langle e_j\, de_i\rangle$, $\langle e_\beta\, de_\alpha\rangle$ $(1 \leqq i, j \leqq 2;\ 3 \leqq \alpha, \beta \leqq 4)$ sind hier genau so wie in 3.1 erklärt.

mit

$$\omega_{ab} = (\delta_{2a}\delta_{1b} - \delta_{2b}\delta_{1a})(\langle e_2 de_1\rangle - \langle e_4 de_3\rangle) + \\ + (\delta_{4a}\delta_{3b} - \delta_{4b}\delta_{3a})(\langle e_2 de_1\rangle + \langle e_4 de_3\rangle) = -\omega_{ba} \qquad (1 \leqq a, b \leqq 4).$$

Daraus folgt:

$$d\omega_{ab} + \sum_{c=1}^{4} \omega_{ac} \wedge \omega_{cb} = \tfrac{1}{2} \sum_{c,d=1}^{4} Q_{abcd}(\omega_c \wedge \omega_d)$$

mit

$$Q_{abcd}(0, \ldots, 0) = R_{abcd}(0, \ldots, 0) = \frac{2}{\lambda_1}(\delta_{2a}\delta_{1b} - \delta_{1a}\delta_{2b})(\delta_{2c}\delta_{1d} - \delta_{1c}\delta_{2d}) + \\ + \frac{2}{\lambda_2}(\delta_{4a}\delta_{3b} - \delta_{3a}\delta_{4b})(\delta_{4c}\delta_{3d} - \delta_{3c}\delta_{4d}) \qquad (1 \leqq a, b, c, d \leqq 4),$$

wobei R_{abcd} die Komponenten des Riemannschen Krümmungstensors der Metrik (7) in bezug auf dasjenige lokale Koordinatensystem (v_1, v_2, v_3, v_4) bedeuten, welches durch die Koordinatentransformation:

$$v_1 = \sqrt{\frac{\lambda_1}{2}}(u_{13} + u_{24})$$

$$v_2 = \sqrt{\frac{\lambda_1}{2}}(u_{23} - u_{14})$$

$$v_3 = \sqrt{\frac{\lambda_2}{2}}(u_{13} - u_{24})$$

$$v_4 = \sqrt{\frac{\lambda_2}{2}}(u_{23} + u_{14})$$

aus den in (1) eingeführten lokalen Koordinaten $(u_{13}, u_{23}, u_{14}, u_{24})$ hervorgeht.

Der Ricci-Tensor R_{ab} berechnet sich nun im Nullpunkt zu:

$$R_{ab}(0, \ldots, 0) = -\frac{2}{\lambda_1}(\delta_{1a}\delta_{1b} + \delta_{2a}\delta_{2b}) - \frac{2}{\lambda_2}(\delta_{3a}\delta_{3b} + \delta_{4a}\delta_{4b}),$$

so daß die Konstanten $\lambda_1/2$ und $\lambda_2/2$ der Ausnahmemetrik geometrisch als die Extremwerte des für beliebige Vektoren $\mathfrak{z} = (z_a) \neq 0$ des Tangentialraums T von $G_{2,2}$ in o erklärten Quotienten

$$-\frac{\sum_{a,b=1}^{4} g_{ab}(0, \ldots, 0)\, z_a z_b}{\sum_{a,b=1}^{4} R_{ab}(0, \ldots, 0)\, z_a z_b} = \frac{(z_1)^2 + (z_2)^2 + (z_3)^2 + (z_4)^2}{\frac{2}{\lambda_1}((z_1)^2 + (z_2)^2) + \frac{2}{\lambda_2}((z_3)^2 + (z_4)^2)}$$

gedeutet werden können. Weiter besitzt die Skalarkrümmung R der Ausnahmemetrik den konstanten Wert $-4\left(\frac{1}{\lambda_1} + \frac{1}{\lambda_2}\right)$, und schließlich findet man für die Riemannsche Krümmung $K(\mathfrak{z}, \mathfrak{w})$ des von den linear unabhängigen Vektoren $\mathfrak{z} = (z_a)$ und $\mathfrak{w} = (w_a)$ aus T aufgespannten Flächenelements:

$$K(\mathfrak{z}, \mathfrak{w}) = 4\,\frac{\frac{1}{\lambda_1}(z_1 w_2 - z_2 w_1)^2 + \frac{1}{\lambda_2}(z_3 w_4 - z_4 w_3)^2}{\sum_{a,b=1}^{4}(z_a w_b - z_b w_a)^2} \geqq 0.$$

K ist nicht konstant, weil sich für $\mathfrak{z}^{(1)}=(\delta_{1a})$, $\mathfrak{w}^{(1)}=(\delta_{2a})$ bzw. $\mathfrak{z}^{(2)}=(\delta_{1a})$, $\mathfrak{w}^{(2)}=(\delta_{3a})$: $K(\mathfrak{z}^{(1)}, \mathfrak{w}^{(1)})=\frac{2}{\lambda_1}>0$ bzw. $K(\mathfrak{z}^{(2)}, \mathfrak{w}^{(2)})=0$ ergibt. Zusammenfassend kann unter Berücksichtigung von (35) festgestellt werden:

SATZ 5. *Die auf der Grassmannschen Mannigfaltigkeit $G_{2,2}$ definierte Riemannsche Metrik* $ds^2=\lambda\left((\langle \mathfrak{e}_3\, d\mathfrak{e}_1\rangle)^2+(\langle \mathfrak{e}_3\, d\mathfrak{e}_2\rangle)^2+(\langle \mathfrak{e}_4\, d\mathfrak{e}_1\rangle)^2+(\langle \mathfrak{e}_4 d\mathfrak{e}_2\rangle)^2\right)+2\mu(\langle \mathfrak{e}_3\, d\mathfrak{e}_1\rangle\langle \mathfrak{e}_4\, d\mathfrak{e}_2\rangle-\langle \mathfrak{e}_3\, d\mathfrak{e}_2\rangle\langle \mathfrak{e}_4\, d\mathfrak{e}_1\rangle)$ $(\lambda, \mu=\text{const},\ \lambda>|\mu|>0)$ *macht $G_{2,2}$ nicht zu einem Einsteinschen Raum. Die Konstanten λ bzw. $|\mu|$ dieser „Ausnahmemetrik" können nämlich geometrisch mit Hilfe des Ricci-Tensors $R_{(i\alpha)(j\beta)}$ als Summe bzw. Absolutbetrag der Differenz der reziproken Extremwerte von*

$$-\frac{\sum_{i,j=1}^{2}\sum_{\alpha,\beta=3}^{4} R_{(i\alpha)(j\beta)}(0,\ldots,0)\,x_{i\alpha}\,x_{j\beta}}{\sum_{i,j=1}^{2}\sum_{\alpha,\beta=3}^{4} g_{(i\alpha)(j\beta)}(0,\ldots,0)\,x_{i\alpha}\,x_{j\beta}}$$ [24]

gedeutet werden, wobei $\mathfrak{x}=(x_{i\alpha})$ einen beliebigen, von Null verschiedenen Vektor des Tangentialraums T von $G_{2,2}$ in o bedeutet. $G_{2,2}$ besitzt bezüglich der Ausnahmemetrik die konstante Skalarkrümmung $R=-\frac{8\lambda}{\lambda^2-\mu^2}$ und eine nichtnegative (aber nicht überall positive) Riemannsche Krümmung; letztere ist nicht konstant.

4. Bemerkungen und Zusätze

4.1. Ein großer Teil der in unseren bisherigen Sätzen aufgeführten Resultate kann auch durch Anwendung der von E. CARTAN stammenden Theorie der „symmetrischen Riemannschen homogenen Räume" erhalten werden[25]). Es handelt sich dabei, kurz gesagt, um folgendes:

Gegeben sei eine zusammenhängende Liesche Gruppe G und eine darin enthaltene kompakte Untergruppe g, deren Lie-Algebra kein von Null verschiedenes Ideal der Lie-Algebra von G enthält. Dann induziert G auf dem (im wesentlichen effektiven) homogenen Raum G/g aller Linksnebenklassen von g (bei Linksmultiplikation) eine transitive Transformationsgruppe, und G/g kann mit (mindestens) einer gegenüber dieser Gruppe invarianten Riemannschen Metrik versehen werden[26]). CARTAN nennt nun den (mit irgendeiner derartigen invarianten Riemannschen Metrik behafteten) Raum G/g einen „*symmetrischen*" Riemannschen homogenen Raum, wenn ein (stetiger) involutorischer Automorphismus σ von G existiert, dessen Fixelemente gerade die Elemente von g sind[27]). Ein derartiger symmetrischer Riemannscher Raum G/g hat die Eigenschaft, daß die kovariante Ableitung seines Riemannschen Krümmungstensors identisch verschwindet, oder — was gleichbedeutend

[24]) Die Tensorkomponenten $g_{(i\alpha)(j\beta)}$ und $R_{(i\alpha)(j\beta)}$ beziehen sich auf die lokalen Koordinaten $u_{i\alpha}$ von $G_{2,2}$.

[25]) Eine zusammenfassende Übersicht über diese Theorie findet sich in der Cartanschen Arbeit: „Les espaces riemanniens symétriques" [7], S. 1247—1255.

[26]) Siehe auch [7], S. 1129—1130.

[27]) Es genügt, hierbei vorauszusetzen, daß die Fixmenge F_σ von σ g enthält und die Komponente der Identität von F_σ in g enthalten ist (die von CARTAN gemachte Voraussetzung, g sei zusammenhängend, kann weitgehend entbehrt werden!). Siehe hierzu [7], S. 1208—1211 und [8], S. 52.

ist — daß eine Parallelverschiebung die Riemannsche Krümmung nicht ändert[28]). Weiter kann im Falle eines (nicht eindimensionalen) kompakten symmetrischen G/g ausgesagt werden, daß seine Riemannsche Krümmung überall nichtnegativ ist[29]). Besonders interessant sind die „*irreduziblen*" symmetrischen Räume G/g, welche sich etwa dadurch charakterisieren lassen, daß sie — von einem konstanten positiven Faktor abgesehen — genau eine gegenüber G invariante Riemannsche Metrik besitzen (vgl. Hilfssatz 1!). Ein symmetrischer Raum G/g ist z.B. irreduzibel, falls G eine einfache Liesche Gruppe ist[30]). Umgekehrt ist bei einem irreduziblen symmetrischen G/g mit halbeinfachem G und zusammenhängendem g entweder G einfach, oder der homogene Raum G/g ist Gruppenraum eines einfachen (abgeschlossenen) Normalteilers von G, dessen Dimension die Hälfte der Dimension von G beträgt[31]). Ein irreduzibler symmetrischer Riemannscher Raum G/g mit zusammenhängendem g ist außerdem stets ein Einsteinscher Raum (vgl. Fußnote 21)[32]). Schließlich gilt: Die (als Lie-Gruppe vorausgesetzte) Komponente der Identität aller lokalen Isometrien eines irreduziblen symmetrischen Riemannschen Raums G/g mit halbeinfachem G besitzt dieselbe Dimension wie G[33]).

Die angeführten Tatsachen können auf unser Problem angewandt werden, da sich die Grassmannschen Mannigfaltigkeiten $G_{l,m}$ bzw. $\widetilde{G_{l,m}}$ als im wesentlichen effektive homogene Räume

$$SO(l+m)/(O(l)\times O(m))\cap SO(l+m) \quad \text{bzw.} \quad SO(l+m)/SO(l)\times SO(m)$$

auffassen lassen, bei welchen durch

$$\sigma\colon \begin{pmatrix} c_{11} & \cdots & c_{1l} & c_{1l+1} & \cdots & c_{1l+m} \\ \vdots & & \vdots & \vdots & & \vdots \\ c_{l1} & \cdots & c_{ll} & c_{ll+1} & \cdots & c_{ll+m} \\ c_{l+11} & \cdots & c_{l+1l} & c_{l+1l+1} & \cdots & c_{l+1l+m} \\ \vdots & & \vdots & \vdots & & \vdots \\ c_{l+m1} & \cdots & c_{l+ml} & c_{l+ml+1} & \cdots & c_{l+ml+m} \end{pmatrix}$$

$$\to \begin{pmatrix} c_{11} & \cdots & c_{1l} & -c_{1l+1} & \cdots & -c_{1l+m} \\ \vdots & & \vdots & \vdots & & \vdots \\ c_{l1} & \cdots & c_{ll} & -c_{ll+1} & \cdots & -c_{ll+m} \\ -c_{l+11} & \cdots & -c_{l+1l} & c_{l+1l+1} & \cdots & c_{l+1l+m} \\ \vdots & & \vdots & \vdots & & \vdots \\ -c_{l+m1} & \cdots & -c_{l+ml} & c_{l+ml+1} & \cdots & c_{l+ml+m} \end{pmatrix}$$

[28]) Vgl. [7], S. 604.

[29]) Vgl. [9], theorem 3.

[30]) Siehe [7], S. 1216. Die von CARTAN gemachte zusätzliche Voraussetzung, G sei kompakt, wird hier nicht benötigt.

[31]) Siehe [7], S. 1216–1217.

[32]) Siehe [7], S. 634.

[33]) Siehe [7], S. 1215.

ein die kompakte Untergruppe $(O(l)\times O(m))\cap SO(l+m)$ (deren Komponente der Identität gleich der kompakten Untergruppe $SO(l)\times SO(m)$ ist) elementweise festlassender, involutorischer Automorphismus von $SO(l+m)$ dargestellt wird. $G_{l,m}$ und $\widetilde{G_{l,m}}$ sind infolgedessen kompakte symmetrische homogene Räume im Sinne CARTANS. Alle gegenüber $SO(n)$ invarianten Riemannschen Metriken dieser Räume haben also (im Fall $lm>1$) eine nichtnegative Riemannsche Krümmung (vgl. Satz 4 und Satz 5!), welche sich bei Parallelverschiebung nicht ändert[34]). Im Falle $l+m\neq 4$ sind wegen der Einfachheit von $SO(n)$ für $n\neq 4$[35]) $G_{l,m}$ und $\widetilde{G_{l,m}}$ irreduzibel und besitzen somit bis auf einen konstanten positiven Faktor nur eine einzige gegen $SO(n)$ invariante Riemannsche Metrik (vgl. Satz 1 und Satz 3!), während die Gruppe $SO(4)$ nicht einfach, ondern nur halbeinfach ist[36]). Daher kann $\widetilde{G_{2,2}}=SO(4)/SO(2)\times SO(2)$ nicht irreduzibel sein, weil sonst $\operatorname{Dim}\widetilde{G_{2,2}}=\frac{1}{2}\operatorname{Dim}SO(4)$ sein müßte (was nicht zutrifft), sodaß also mehr als eine nur von einem konstanten positiven Faktor abhängige, gegen $SO(4)$ invariante Riemannsche Metriken von $\widetilde{G_{2,2}}$ existieren müssen (vgl. Satz 3!). Aus der schon gezeigten Irreduzibilität von $\widetilde{G_{l,m}}$ im Falle $l+m\neq 4$ folgt weiter die Tatsache, daß die gegen $SO(n)$ invarianten Riemannschen Metriken $\widetilde{G_{l,m}}$ zu einem Einsteinschen Raum machen (vgl. Satz 4!)[37]). Schließlich ist die Komponente der Identität aller lokalen Isometrien eines $G_{l,m}$ bzw. $\widetilde{G_{l,m}}$ mit $lm>1$ von konstanter Riemannscher Krümmung $\frac{1}{2}lm(lm+1)$-dimensional[38]) und kann aus diesem Grunde nur dann dieselbe Dimension wie $SO(n)$ haben, wenn $\frac{1}{2}lm(lm+1)=\frac{1}{2}(l+m-1)(l+m)$, d.h. $l=1$ oder $m=1$ gilt. $G_{l,m}$ bzw. $\widetilde{G_{l,m}}$ mit $lm>1$ und $l+m\neq 4$ besitzen daher nur dann konstante Riemannsche Krümmung in bezug auf die gegen $SO(n)$ invarianten Riemannschen Metriken, wenn $l=1$ oder $m=1$ ist (vgl. Satz 4!).

4.2. Man kann analog zu den reellen Grassmannschen Mannigfaltigkeiten $G_{l,m}$ und $\widetilde{G_{l,m}}$ auch die *komplexe* Grassmannsche Mannigfaltigkeit $H_{l,m}$ als Menge aller (komplexen) l-dimensionalen Untervektorräume des komplexen n-dimensionalen Vektorraums C_n definieren ($n=l+m$). $H_{l,m}$ erweist sich dann als eine komplex analytische Mannigfaltigkeit der reellen Dimension $2lm$, auf welcher analog wie bei (1) lokale komplexe Koordinaten $w_{i\alpha}$ ($i=1,\ldots,l$; $\alpha=l+1,\ldots,n$) eingeführt werden können. Während wir früher in $G_{l,m}$ und $\widetilde{G_{l,m}}$ gegen $SO(n)$ bzw. $O(n)$ invariante Riemannsche Metriken aufsuchten, sollen jetzt entsprechend dazu *Hermitesche* Metriken von $H_{l,m}$ betrachtet werden, welche gegenüber der auf $H_{l,m}$ wirkenden *unitären unimodularen* Gruppe $SU(n)$ bzw. der *unitären* Gruppe $U(n)$ invariant sind. Dabei wird

---

[34]) Dies läßt sich für unsere durch (6) und (7) gegebenen Riemannschen Metriken auch leicht direkt durch Nachrechnung mit Hilfe des Cartanschen Kalküls bestätigen.

[35]) Siehe [*10*], theorem 82.

[36]) Siehe [*10*], S. 273, ex. 70.

[37]) Dasselbe gilt auch für $G_{l,m}$, wie eine genauere Analyse der Cartanschen Sätze zeigt.

[38]) Siehe [*6*], S. 88.

unter $U(n)$ die Gruppe aller linearen Transformationen des C_n verstanden, welche eine (ausgezeichnete) Hermitesche Bilinearform $\langle \bar{\mathfrak{x}}\, \mathfrak{y}\rangle$[39] $(\mathfrak{x}, \mathfrak{y} \in C_n)$ mit positiv definiter Hermitescher Form $\langle \bar{\mathfrak{x}}\, \mathfrak{x}\rangle$ invariant lassen, während die Transformationen von $SU(n)$ außerdem noch die Determinante 1 besitzen.

Denkt man sich nun die das lokale Koordinatensystem $\{w_{i\alpha}\}$ bestimmenden Basisvektoren $\mathring{\mathfrak{e}}_i$ und $\mathring{\mathfrak{e}}_\alpha$ unitärnormiert: $\langle \bar{\mathring{\mathfrak{e}}}_i \mathring{\mathfrak{e}}_j\rangle = \delta_{ij}$, $\langle \bar{\mathring{\mathfrak{e}}}_\alpha \mathring{\mathfrak{e}}_\beta\rangle = \delta_{\alpha\beta}$, $\langle \bar{\mathring{\mathfrak{e}}}_i \mathring{\mathfrak{e}}_\alpha\rangle = 0$ $(1 \leqq i, j \leqq l;\ l+1 \leqq \alpha, \beta \leqq n)$, so induziert die den (durch $\{\mathring{\mathfrak{e}}_i\}$ bestimmten) Punkt $o \in H_{l,m}$ festlassende Untergruppe von $SU(n)$ im Tangentialraum U von $H_{l,m}$ in o die durch:

$$z'_{i\alpha} = \sum_{j=1}^{l} \sum_{\beta=l+1}^{n} \bar{d}_{ij}\, e_{\alpha\beta}\, z_{j\beta}$$

($z_{i\alpha}$ = Komponenten der Vektoren von U; $i = 1, \ldots, l$; $\alpha = l+1, \ldots, n$)

mit

$$(d_{ij}) \in U(l), \qquad (e_{\alpha\beta}) \in U(m), \qquad \|d_{ij}\|\,\|e_{\alpha\beta}\| = 1$$

gegebene lineare Isotropiegruppe $\big(U(l) \otimes U(m)\big)^*$ (vgl. (4)!). Letztere ist in $U(n)$ enthalten; und U ist in bezug auf diese lineare Isotropiegruppe irreduzibel[40], weil aus der Irreduzibilität von C_l bezüglich $SU(l)$ und von C_m bezüglich $SU(m)$ ($SU(l)$ bzw. $SU(m)$ operiert transitiv auf den „Richtungen" von C_l bzw. C_m!) nach einem Satz bei H. WEYL[41] auch die Irreduzibilität von U bezüglich der in $\big(U(l) \otimes U(m)\big)^*$ enthaltenen Untergruppe $SU(l) \otimes SU(m)$ folgt[42]. Infolgedessen ergeben ähnlich wie beim Beweis von Hilfssatz 1 durchgeführte Überlegungen, daß die Hermiteschen Formen $\lambda \left(\sum_{i=1}^{l} \sum_{\alpha=l+1}^{n} \bar{\mathfrak{x}}_{i\alpha}\, z_{i\alpha} \right)$ mit konstantem positivem λ die einzigen, gegenüber $\big(U(l) \otimes U(m)\big)^*$ invarianten, positiv definiten Hermiteschen Formen auf U sind. Daraus folgt aber schon:

SATZ 6. *Die durch*

$$ds^2 = \sum_{i=1}^{l} \sum_{\alpha=l+1}^{n} |\langle \bar{\mathfrak{x}}_\alpha\, d\mathfrak{e}_i\rangle|^2 \tag{36}$$

gegebene Hermitesche Metrik ist, von einem konstanten positiven Proportionalitätsfaktor λ abgesehen, die einzige gegenüber sowohl allen Transformationen von $SU(n)$ als auch allen Transformationen von $U(n)$ invariante Hermitesche Metrik auf der komplexen Grassmannschen Mannigfaltigkeit $H_{l,m}$[43].

---

[39] Unter einer Hermiteschen Bilinearform $\langle \bar{\mathfrak{x}}\, \mathfrak{y}\rangle$ wird ein in $\bar{\mathfrak{x}}$ und $\mathfrak{y}$ lineares (komplexwertiges) Funktional mit der Symmetrieeigenschaft $\langle \bar{\mathfrak{y}}\, \mathfrak{x}\rangle = \overline{\langle \bar{\mathfrak{x}}\, \mathfrak{y}\rangle}$ verstanden. Vgl. dazu [*11*], S. 15–18.

[40] Die Irreduzibilität eines komplexen Vektorraums in bezug auf eine darauf erklärte lineare Gruppe wird analog wie für reelle Vektorräume (vgl. Def. 1 und 2!) definiert.

[41] Siehe [*11*], S. 126–127.

[42] Vgl. [*12*], S. 407.

[43] Hierbei beziehen sich die Differentialformen auf ein von den $2lm$ lokalen, reellen Koordinaten $\operatorname{Re} w_{i\alpha}$ und $\operatorname{Im} w_{i\alpha}$ von $H_{l,m}$ beliebig oft stetig differenzierbar abhängendes, unitärnormiertes Vektorfeld $\{\mathfrak{e}_j, \mathfrak{e}_\beta\}$ mit der Nebenbedingung $\mathfrak{e}_j(0, \ldots, 0) = \mathring{\mathfrak{e}}_j$, $\mathfrak{e}_\beta(0, \ldots, 0) = \mathring{\mathfrak{e}}_\beta$ $(i, j = 1, \ldots, 1;\ \alpha, \beta = l+1, \ldots, n)$ und derart, daß jeder Punkt von $H_{l,m}$ mit den Koordinaten $w_{i\alpha}$ durch das l-Bein $\{\mathfrak{e}_j(\operatorname{Re} w_{i\alpha}, \operatorname{Im} w_{i\alpha})\}$ dargestellt wird.

Wir bemerken noch: Faßt man $H_{l,m}$ als $2lm$-dimensionale reell analytische Mannigfaltigkeit auf, so erweist sich die Hermitesche Metrik (36) nach der Cartanschen Theorie wegen der Einfachheit der Lieschen Gruppe $SU(l+m)$[44]) des symmetrischen homogenen Raums $H_{l,m}=SU(l+m)/(U(l)\times U(m))\cap SU(l+m)$ sogar als die — von einem konstanten positiven Proportionalitätsfaktor abgesehen — einzige gegen $SU(n)$ invariante *Riemannsche* Metrik von $H_{l,m}$ (vgl. 4.1!). Außerdem soll vermerkt werden, *daß die Hermitesche Metrik* (36) *eine Kählersche Metrik darstellt*[45]); es gilt nämlich mit den (23), (30) und (31) entsprechenden Abkürzungen $\omega_{(i\alpha)}=\langle\bar{e}_\alpha\, de_i\rangle=-\overline{\omega_{(\alpha i)}}$, $\omega_{(ij)}=\langle\bar{e}_j\, de_i\rangle=-\overline{\omega_{(ji)}}$, $\omega_{(\alpha\beta)}=\langle\bar{e}_\beta\, de_\alpha\rangle=-\overline{\omega_{(\beta\alpha)}}$ wegen

$$(37)\qquad \begin{cases} de_i=\sum_{j=1}^{l}\omega_{(ij)}e_j+\sum_{\beta=l+1}^{n}\omega_{(i\beta)}e_\beta \\ de_\alpha=\sum_{j=1}^{l}\omega_{(\alpha j)}e_j+\sum_{\beta=l+1}^{n}\omega_{(\alpha\beta)}e_\beta \qquad (i=1,\ldots,l;\ \alpha=l+1,\ldots,n) \end{cases}$$

für die Differentialform $\omega=\sum_{i=1}^{l}\sum_{\alpha=l+1}^{n}\overline{\omega_{(i\alpha)}}\wedge\omega_{(i\alpha)}$:

$$\begin{aligned} d\omega&=\sum_{i=1}^{l}\sum_{\alpha=l+1}^{n}\left(d\,\overline{\omega_{(i\alpha)}}\wedge\omega_{(i\alpha)}-\overline{\omega_{(i\alpha)}}\wedge d\omega_{(i\alpha)}\right)\\ &=\sum_{i=1}^{l}\sum_{\alpha=l+1}^{n}\Big(\Big(\sum_{j=1}^{l}\omega_{(\alpha j)}\wedge\overline{\omega_{(ij)}}+\sum_{\beta=l+1}^{n}\omega_{(\alpha\beta)}\wedge\overline{\omega_{(i\beta)}}\Big)\wedge\omega_{(i\alpha)}-\\ &\quad-\overline{\omega_{(i\alpha)}}\wedge\Big(\sum_{j=1}^{l}\overline{\omega_{(\alpha j)}}\wedge\omega_{(ij)}+\sum_{\beta=l+1}^{n}\overline{\omega_{(\alpha\beta)}}\wedge\omega_{(i\beta)}\Big)\Big)\\ &=\sum_{i,j=1}^{l}\sum_{\alpha=l+1}^{n}\left(\omega_{(\alpha i)}\wedge\overline{\omega_{(ji)}}\wedge\omega_{(j\alpha)}+\omega_{(\alpha i)}\wedge\omega_{(j\alpha)}\wedge\overline{\omega_{(ji)}}+\right.\\ &\quad\left.+\sum_{i=1}^{l}\sum_{\alpha,\beta=l+1}^{n}\left(\omega_{(\beta\alpha)}\wedge\overline{\omega_{(i\alpha)}}\wedge\omega_{(i\beta)}+\overline{\omega_{(i\alpha)}}\wedge\omega_{(\beta\alpha)}\wedge\omega_{(i\beta)}\right)=0.\right. \end{aligned}$$

Diese Kählersche Metrik (36) stimmt im Fall $l=1$ wegen

$$(38)\qquad ds^2=\sum_{\alpha=2}^{n}|\langle\bar{e}_\alpha\, de_1\rangle|^2=\langle d\bar{e}_1\, de_1\rangle-\langle e_1\, d\bar{e}_1\rangle\langle\bar{e}_1\, de_1\rangle$$

mit der als „elliptisch" bezeichneten Hermiteschen Metrik des $(n-1)$-dimensionalen komplexen projektiven Raums überein[46]).

Schließlich führen analog wie in Abschnitt 3 der Arbeit durchgeführte Rechnungen zu folgenden Resultaten für die Krümmungseigenschaften des als $2lm$-dimensionale reell analytische Mannigfaltigkeit aufgefaßten und mit der Riemannschen Metrik $ds^2=\lambda\Big(\sum_{i=1}^{l}\sum_{\alpha=l+1}^{n}|\langle\bar{e}_\alpha\, de_i\rangle|^2\Big)$ $(\lambda>0)$ versehenen Raums $H_{l,m}$: Die Komponenten des Riemannschen Krümmungstensors bezüglich der

[44]) Siehe [*10*], theorem 82.
[45]) Vgl. [*13*].
[46]) Siehe [*14*], S. 116.

lokalen Koordinaten $\mathrm{Re}\, w_{i\alpha}$ und $\mathrm{Im}\, w_{i\alpha}$ [vgl. Fußnote [43])] von $H_{l,m}$ im Nullpunkt berechnen sich zu

$$R_{(i\alpha)(j\beta)(k\gamma)(h\delta)}(0,\ldots,0) = R_{(i\alpha)(j\beta)(\overline{k\gamma})(\overline{h\delta})}(0,\ldots,0)$$
$$= R_{(\overline{i\alpha})(\overline{j\beta})(k\gamma)(h\delta)}(0,\ldots,0) = R_{(\overline{i\alpha})(\overline{j\beta})(\overline{k\gamma})(\overline{h\delta})}(0,\ldots,0)$$
$$= \lambda\left((\delta_{ik}\,\delta_{jh} - \delta_{ih}\,\delta_{jk})\,\delta_{\alpha\beta}\,\delta_{\gamma\delta} + \delta_{ij}\,\delta_{kh}(\delta_{\alpha\gamma}\,\delta_{\beta\delta} - \delta_{\alpha\delta}\,\delta_{\beta\gamma})\right)$$

und

$$R_{(i\alpha)(\overline{j\beta})(k\gamma)(\overline{h\delta})}(0,\ldots,0) = -\,R_{(i\alpha)(\overline{j\beta})(\overline{k\gamma})(h\delta)}(0,\ldots,0)$$
$$= R_{(\overline{i\alpha})(j\beta)(\overline{k\gamma})(h\delta)}(0,\ldots,0) = -\,R_{(\overline{i\alpha})(j\beta)(k\gamma)(\overline{h\delta})}(0,\ldots,0)$$
$$= \lambda\left((\delta_{ik}\,\delta_{jh} + \delta_{ih}\,\delta_{jk})\,\delta_{\alpha\beta}\,\delta_{\gamma\delta} + \delta_{ij}\,\delta_{kh}(\delta_{\alpha\gamma}\,\delta_{\beta\delta} + \delta_{\alpha\delta}\,\delta_{\beta\gamma})\right)$$ [47]),

während die restlichen Komponenten des Krümmungstensors gleich Null sind. *Daraus folgt, daß $H_{l,m}$ einen Einsteinschen Raum mit der konstanten Skalarkrümmung $R = -\frac{4}{\lambda}\,l\,m\,(l+m)$ darstellt, dessen Riemannsche Krümmung K zwar überall nichtnegativ, aber nur dann überall positiv ist, wenn $l=1$ oder $m=1$ gilt. In diesem Fall ($H_{l,m}$ = komplexer projektiver Raum) besitzt K nur für $l=m=1$ den konstanten Wert $4/\lambda$, während K sonst alle Werte zwischen $1/\lambda$ und $4/\lambda$ annimmt. Auch in den übrigen Fällen $l>1$ und $m>1$ ist K nicht konstant.*

4.3. Bildet man jeden l-dimensionalen Untervektorraum V_l des n-dimensionalen euklidischen Vektorraums V_n auf den zu V_l bezüglich V_n total orthogonalen m-dimensionalen Untervektorraum V_m ab, so entsteht eine homöomorphe Abbildung A von $G_{l,m}$ auf $G_{m,l}$. Analog dazu kann ein Homöomorphismus $\tilde{A}$ von $\widetilde{G_{l,m}}$ auf $\widetilde{G_{m,l}}$ sowie ein Homöomorphismus B von $H_{l,m}$ auf $H_{m,l}$ erklärt werden. Alle diese Abbildungen sind (beliebig oft) differenzierbar, wie man etwa im Falle von B folgendermaßen einsieht: Die Vektoren $\mathfrak{f}_i = \mathring{\mathfrak{e}}_i + \sum_{\beta=l+1}^{n} w_{i\beta}\,\mathring{\mathfrak{e}}_\beta$ $(i=1,\ldots,l)$ sind zu den Vektoren $\mathfrak{g}_\alpha = \mathring{\mathfrak{e}}_\alpha - \sum_{j=1}^{l} \overline{w_{j\alpha}}\,\mathring{\mathfrak{e}}_j$ $(\alpha=l+1,\ldots,n)$ „orthogonal", d.h. es gilt $\bar{\mathfrak{f}}_i\,\mathfrak{g}_\alpha = 0$ (wenn $\{\mathring{\mathfrak{e}}_i, \mathring{\mathfrak{e}}_\alpha\}$ unitärnormiert gewählt wird), so daß sich B in den durch die n-Beine $\{\mathring{\mathfrak{e}}_i, \mathring{\mathfrak{e}}_\alpha\}$ bzw. $\{\mathring{\mathfrak{e}}_\alpha, \mathring{\mathfrak{e}}_i\}$ bestimmten, lokalen Koordinaten $\mathrm{Re}\, w_{i\alpha}$ und $\mathrm{Im}\, w_{i\alpha}$ bzw. $\mathrm{Re}\, w_{\alpha i}$ und $\mathrm{Im}\, w_{\alpha i}$ von $H_{l,m}$ bzw. $H_{m,l}$ durch:

$$\left.\begin{aligned}\mathrm{Re}\, w_{\alpha i} &= -\,\mathrm{Re}\, w_{i\alpha}\\ \mathrm{Im}\, w_{\alpha i} &= \mathrm{Im}\, w_{i\alpha}\end{aligned}\right\} \qquad (i=1,\ldots,l;\ \alpha=l+1,\ldots,n)$$

darstellt.

Jetzt gilt: *A und $\tilde{A}$ bzw. B sind isometrische Abbildungen der mit den Normalmetriken $\lambda\left(\sum_{i=1}^{l}\sum_{\alpha=l+1}^{n}\langle \mathfrak{e}_\alpha\, d\,\mathfrak{e}_i\rangle^2\right)$ und $\lambda\left(\sum_{\alpha=l+1}^{n}\sum_{i=1}^{l}\langle \mathfrak{e}_i\, d\,\mathfrak{e}_\alpha\rangle^2\right)$ versehenen Räume*

[47]) Hierbei beziehen sich die nichtüberstrichenen bzw. die überstrichenen Komponenten des Krümmungstensors auf die Koordinaten $\mathrm{Re}\, w_{i\alpha}$ bzw. $\mathrm{Im}\, w_{i\alpha}$ von $H_{l,m}$.

$G_{l,m}$ und $G_{m,l}$ sowie (im Falle $l=m=2$) des mit den Ausnahmemetriken

$$\lambda(\langle e_3\,de_1\rangle^2+\langle e_3\,de_2\rangle^2+\langle e_4\,de_1\rangle^2+\langle e_4\,de_2\rangle^2)+ \\ +2\mu(\langle e_3\,de_1\rangle\langle e_4\,de_2\rangle-\langle e_3\,de_2\rangle\langle e_4\,de_1\rangle)$$

und

$$\lambda(\langle e_1\,de_3\rangle^2+\langle e_1\,de_4\rangle^2+\langle e_2\,de_3\rangle^2+\langle e_2\,de_4\rangle^2)+ \\ +2\mu(\langle e_1\,de_3\rangle\langle e_2\,de_4\rangle-\langle e_1\,de_4\rangle\langle e_2\,de_3\rangle)$$

versehenen Raums $G_{2,2}$ bzw. der mit den Hermiteschen Metriken $\lambda\left(\sum_{i=1}^{l}\sum_{\alpha=l+1}^{n}|\langle\bar{e}_\alpha\,de_i\rangle|^2\right)$ und $\lambda\left(\sum_{\alpha=l+1}^{n}\sum_{i=1}^{l}|\langle\bar{e}_i\,de_\alpha\rangle|^2\right)$ versehenen Räume $H_{l,m}$ und $H_{m,l}$. Dies folgt nämlich unmittelbar aus den Beziehungen $\langle e_\alpha\,de_i\rangle=-\langle e_i\,de_\alpha\rangle$ bzw. $\langle\bar{e}_\alpha\,de_i\rangle=-\overline{\langle\bar{e}_i\,de_\alpha\rangle}$ $(i=1,\ldots,l;\ \alpha=l+1,\ldots,n)$.

4.4. Sind $\mathfrak{v}_1,\ldots,\mathfrak{v}_l$ l Vektoren eines (reellen oder komplexen) n-dimensionalen Vektorraums $(1\leqq l\leqq n-1)$, so werde im folgenden mit $[\mathfrak{v}_1\ldots\mathfrak{v}_l]$ derjenige Vektor des $\binom{n}{l}$-dimensionalen (reellen oder komplexen) Vektorraums bezeichnet, dessen Komponenten die (irgendwie angeordneten) l-reihigen Unterdeterminanten der aus den (sich auf eine beliebige Basis des n-dimensionalen Vektorraums beziehenden) Komponenten $\alpha_i^{(\nu)}$ von $\mathfrak{v}_i$ gebildeten Matrix $(\alpha_i^{(\nu)})$ sind $(i=1,\ldots,l;\ \nu=1,\ldots,n)$. Nun können bekanntlich die Grassmannschen Mannigfaltigkeiten $G_{l,m}$ bzw. $H_{l,m}$ als algebraische Untermannigfaltigkeiten des $\left(\binom{n}{l}-1\right)$-dimensionalen reellen bzw. komplexen projektiven Raums aufgefaßt werden[48]), indem man jedem Punkt p aus $G_{l,m}$ bzw. $H_{l,m}$ als „Grassmannsche projektive Koordinaten" die (nicht alle verschwindenden) Komponenten von $[\mathfrak{v}_1,\ldots,\mathfrak{v}_l]$ zuordnet, wobei $\mathfrak{v}_1,\ldots,\mathfrak{v}_l$ l linear unabhängige, sonst aber beliebige Vektoren aus dem p bestimmenden l-dimensionalen Untervektorraum des n-dimensionalen reellen bzw. komplexen Vektorraums V_n bzw. C_n sind. Bezieht man die Vektoren $\mathfrak{v}_1,\ldots,\mathfrak{v}_l$ und $\mathfrak{w}_1,\ldots,\mathfrak{w}_l$ dieses n-dimensionalen Vektorraums auf eine orthonormierte bzw. unitärnormierte Basis $\{\mathring{e}_\nu\}$ $(\nu=1,\ldots,n)$ desselben und führt man in dem $\binom{n}{l}$-dimensionalen Vektorraum in trivialer Weise ein inneres Produkt bzw. eine Hermitesche Bilinearform ein, so gilt auf Grund des verallgemeinerten Determinantenmultiplikationssatzes:

$$\langle[\mathfrak{v}_1\ldots\mathfrak{v}_l]\,[\mathfrak{w}_1\ldots\mathfrak{w}_l]\rangle=\|\langle\mathfrak{v}_i\,\mathfrak{w}_j\rangle\| \tag{39}$$

bzw.

$$\langle[\bar{\mathfrak{v}}_1\ldots\bar{\mathfrak{v}}_l]\,[\mathfrak{w}_1\ldots\mathfrak{w}_l]\rangle=\|\langle\bar{\mathfrak{v}}_i\,\mathfrak{w}_j\rangle\| \qquad (i,j=1,\ldots,l). \tag{40}$$

Jetzt ist es nicht schwer einzusehen, daß die beschriebene „Plückersche Einbettung" von $G_{l,m}$ bzw. $H_{l,m}$ in den $\left(\binom{n}{l}-1\right)$-dimensionalen reellen bzw.

[48]) Siehe [15]. S. 312.

komplexen projektiven Raum $P_{\binom{n}{l}-1}$ bzw. $Q_{\binom{n}{l}-1}$ durch eine „reguläre" Abbildung vermittelt wird. Darunter verstehen wir eine Abbildung, die sich in den lokalen reellen bzw. komplexen Koordinaten $u_{i\alpha}$ bzw. $w_{i\alpha}$ von $G_{l,m}$ bzw. $H_{l,m}$ und den lokalen reellen bzw. komplexen Koordinaten u_I bzw. w_I von $P_{\binom{n}{l}-1}=G_{1,\binom{n}{l}-1}$ bzw. $Q_{\binom{n}{l}-1}=H_{1,\binom{n}{l}-1}$ durch

$$u_I=\Phi_I(u_{i\alpha}) \quad \text{bzw.} \quad w_I=\Psi_I(w_{i\alpha})$$

mit (beliebig oft) stetig nach allen Variablen differenzierbaren Funktionen Φ_I bzw. Ψ_I ausdrücken läßt, deren Funktionalmatrizen $\left(\frac{\partial \Phi_I}{\partial u_{i\alpha}}\right)$ bzw. $\left(\frac{\partial \Psi_I}{\partial w_{i\alpha}}\right)$ überall Maximalrang besitzen (hierbei durchläuft der Index I etwa alle Kombinationen $(\nu_1,\ldots,\nu_l)$ mit $1\leqq\nu_1<\nu_2<\cdots<\nu_l\leqq n$ und $\nu_l>l$). Legt man nämlich die lokalen Koordinaten von $G_{l,m}$ bzw. $H_{l,m}$ mittels des n-Beins $\{\mathfrak{e}_\nu\}$ und legt man die lokalen Koordinaten von $G_{1,\binom{n}{l}-1}$ bzw. $H_{1,\binom{n}{l}-1}$ mittels des wegen (39) und (40) orthonormierten bzw. unitärnormierten $\binom{n}{l}$-Beins $\{[\mathfrak{e}_{\nu_1}\ldots\mathfrak{e}_{\nu_l}]\}$ $(1\leqq\nu_1<\cdots<\nu_l\leqq n)$ fest, so wird jeder in einer Umgebung des Punkts $u_I=0$ bzw. $w_I=0$ gelegene Punkt aus $G_{1,\binom{n}{l}-1}$ bzw. $H_{1,\binom{n}{l}-1}$ durch einen Vektor

$$(41)\quad \begin{cases} \mathfrak{f}_{1\ldots l}=[\mathfrak{e}_1\ldots\mathfrak{e}_l]+\sum\limits_{1\leqq\nu_1<\cdots<\nu_l\leqq n,\ l<\nu_l} u_{\nu_1\ldots\nu_l}[\mathfrak{e}_{\nu_1}\ldots\mathfrak{e}_{\nu_l}] \\ \text{bzw.} \\ \mathfrak{f}_{1\ldots l}=[\mathfrak{e}_1\ldots\mathfrak{e}_l]+\sum\limits_{1\leqq\nu_1<\cdots<\nu_l\leqq n,\ l<\nu_l} w_{\nu_1\ldots\nu_l}[\mathfrak{e}_{\nu_1}\ldots\mathfrak{e}_{\nu_l}] \end{cases}$$

dargestellt. Weiter läßt sich jeder in einer Umgebung von $w_{i\alpha}=0$ gelegene Punkt aus $H_{l,m}$ durch die Vektoren $\mathfrak{f}_i=\mathfrak{e}_i+\sum\limits_{\alpha=l+1}^{n} w_{i\alpha}\mathfrak{e}_\alpha$ $(i=1,\ldots,l)$ darstellen und besitzt somit als Grassmannsche Koordinaten die Komponenten des Vektors $[\mathfrak{f}_1\ldots\mathfrak{f}_l]$, welcher sich folgendermaßen durch die Basisvektoren $[\mathfrak{e}_{\nu_1}\ldots\mathfrak{e}_{\nu_l}]$ linear kombinieren läßt:

$$(42)\quad \begin{cases} [\mathfrak{f}_1\ldots\mathfrak{f}_l]=[\mathfrak{e}_1\ldots\mathfrak{e}_l]+\sum\limits_{i=1}^{l}\sum\limits_{\alpha=l+1}^{n}(-1)^{l-i}w_{i\alpha}[\mathfrak{e}_1\ldots\mathfrak{e}_{i-1}\mathfrak{e}_{i+1}\ldots\mathfrak{e}_l\mathfrak{e}_\alpha]+ \\ \qquad+\sum\limits_{1\leqq\nu_1<\cdots<\nu_l\leqq n,\ l<\nu_{l-1}} F_{\nu_1\ldots\nu_l}(w_{j\beta})[\mathfrak{e}_{\nu_1}\ldots\mathfrak{e}_{\nu_l}] \\ \qquad\qquad (1\leqq j\leqq l;\ l+1\leqq\beta\leqq n), \end{cases}$$

wobei die $F_{\nu_1\ldots\nu_l}$ homogene Polynome in den $w_{j\beta}$ von mindestens zweitem Grade sind. Durch Vergleich von (41) und (42) folgt jetzt für die Funktionen Ψ_I unserer Abbildung von $H_{l,m}$ in $H_{1,\binom{n}{l}-1}$:

$$\Psi_{1\ldots i-1\,i+1\ldots l\,\alpha}(w_{j\beta})=(-1)^{l-i}w_{i\alpha} \qquad (1\leqq i\leqq l;\ l+1\leqq\alpha\leqq n)$$

$$\Psi_{\nu_1\ldots\nu_l}(w_{j\beta})=F_{\nu_1\ldots\nu_l}(w_{j\beta}) \qquad (1\leqq\nu_1<\cdots<\nu_l\leqq n;\ l<\nu_{l-1}),$$

und analoge Formeln gelten auch für die Funktionen Φ_I. Die Funktionen Φ_I und Ψ_I besitzen also in der Tat die angegebenen Eigenschaften und vermitteln somit eine reguläre Abbildung. Nun kann man außerdem zeigen:

SATZ 7. *Die bei der Plückerschen Einbettung der Grassmannschen Mannigfaltigkeiten $G_{l,m}$ bzw. $H_{l,m}$ in die projektiven Räume $P_{\binom{n}{l}-1}$ bzw. $Q_{\binom{n}{l}-1}$ durch die elliptischen Riemannschen bzw. elliptischen Hermiteschen Metriken $dS^2 = \lambda(\langle d\mathfrak{e}_{1\dots l}\, d\mathfrak{e}_{1\dots l}\rangle)$ bzw. $dS^2 = \lambda\big(\langle d\overline{\mathfrak{e}_{1\dots l}}\, d\mathfrak{e}_{1\dots l}\rangle - \langle \mathfrak{e}_{1\dots l}\, d\overline{\mathfrak{e}_{1\dots l}}\rangle\langle \overline{\mathfrak{e}_{1\dots l}}\, d\mathfrak{e}_{1\dots l}\rangle\big)$ $(\lambda>0)$ (vgl. Fußnote [8]) und (38)) von $P_{\binom{n}{l}-1}$ bzw. $Q_{\binom{n}{l}-1}$ auf $G_{l,m}$ bzw. $H_{l,m}$ induzierten Riemannschen bzw. Hermiteschen Metriken stimmen mit den gegen $O(n)$ bzw. $U(n)$ invarianten Metriken $ds^2 = \lambda\Big(\sum_{i=1}^{l}\sum_{\alpha=l+1}^{n}\langle \mathfrak{e}_\alpha\, d\mathfrak{e}_i\rangle^2\Big)$ bzw. $ds^2 = \lambda\Big(\sum_{i=1}^{l}\sum_{\alpha=l+1}^{n}|\langle \overline{\mathfrak{e}}_\alpha\, d\mathfrak{e}_i\rangle|^2\Big)$ $(\lambda>0)$ von $G_{l,m}$ bzw. $H_{l,m}$ überein.*

BEWEIS. Wir wollen den Beweis der die Einbettung von $H_{l,m}$ in $Q_{\binom{n}{l}-1}$ betreffenden Behauptung von Satz 7 durchführen; der Beweis der Behauptung über die Einbettung von $G_{l,m}$ in $P_{\binom{n}{l}-1}$ verläuft dann ganz analog. Zu diesem Zweck benutzen wir, daß die durch die angegebene elliptische Hermitesche Metrik von $Q_{\binom{n}{l}-1}$ in $H_{l,m}$ induzierte Hermitesche Metrik das Linienelementquadrat

$$ds_1^2 = \lambda\big(\langle d[\overline{\mathfrak{e}}_1 \dots \overline{\mathfrak{e}}_l]\, d[\mathfrak{e}_1 \dots \mathfrak{e}_l]\rangle - \langle [\mathfrak{e}_1 \dots \mathfrak{e}_l]\, d[\overline{\mathfrak{e}}_1 \dots \overline{\mathfrak{e}}_l]\rangle\langle [\overline{\mathfrak{e}}_1 \dots \overline{\mathfrak{e}}_l]\, d[\mathfrak{e}_1 \dots \mathfrak{e}_l]\rangle\big)$$

besitzt, wobei die $\mathfrak{e}_j = \mathfrak{e}_j(\operatorname{Re} w_{i\alpha}, \operatorname{Im} w_{i\alpha})$ solche (beliebig oft) stetig differenzierbar von den lokalen reellen Koordinaten von $H_{l,m}$ abhängende, unitärnormierte Vektorfunktionen mit der Nebenbedingung $\mathfrak{e}_j(0, \dots, 0) = \mathring{\mathfrak{e}}_j$ sind $(1 \leqq i, j \leqq l;\ l+1 \leqq \alpha \leqq n)$, daß der Punkt mit den Koordinaten $w_{i\alpha}$ von $H_{l,m}$ durch die Vektoren $\mathfrak{e}_j(\operatorname{Re} w_{i\alpha}, \operatorname{Im} w_{i\alpha})$ dargestellt wird. Damit erhalten wir zusammen mit (37) und (40):

$$\begin{aligned}
ds_1^2 &= \lambda \sum_{i,j=1}^{l} \big(\langle [\overline{\mathfrak{e}}_1 \dots \overline{\mathfrak{e}}_{i-1}\, d\overline{\mathfrak{e}}_i\, \overline{\mathfrak{e}}_{i+1} \dots \overline{\mathfrak{e}}_l]\, [\mathfrak{e}_1 \dots \mathfrak{e}_{j-1}\, d\mathfrak{e}_j\, \mathfrak{e}_{j+1} \dots \mathfrak{e}_l]\rangle - \\
&\quad - \langle [\mathfrak{e}_1 \dots \mathfrak{e}_l]\, [\overline{\mathfrak{e}}_1 \dots \overline{\mathfrak{e}}_{i-1}\, d\overline{\mathfrak{e}}_i\, \overline{\mathfrak{e}}_{i+1} \dots \overline{\mathfrak{e}}_l]\rangle \langle [\overline{\mathfrak{e}}_1 \dots \overline{\mathfrak{e}}_l]\, [\mathfrak{e}_1 \dots \mathfrak{e}_{j-1}\, d\mathfrak{e}_j\, \mathfrak{e}_{j+1} \dots \mathfrak{e}_l]\rangle\big) \\
&= \lambda \sum_{i,j=1}^{l} \Big\langle \Big(\overline{\omega_{ii}}\, [\overline{\mathfrak{e}}_1 \dots \overline{\mathfrak{e}}_l] + \sum_{\alpha=l+1}^{n} \overline{\omega_{i\alpha}}\, [\overline{\mathfrak{e}}_1 \dots \overline{\mathfrak{e}}_{i-1}\, \overline{\mathfrak{e}}_\alpha\, \overline{\mathfrak{e}}_{i+1} \dots \overline{\mathfrak{e}}_l] - \overline{\omega_{ii}}\, [\overline{\mathfrak{e}}_1 \dots \overline{\mathfrak{e}}_l]\Big) \times \\
&\quad \times [\mathfrak{e}_1 \dots \mathfrak{e}_{j-1}\, d\mathfrak{e}_j\, \mathfrak{e}_{j+1} \dots \mathfrak{e}_l]\Big\rangle \\
&= \lambda \sum_{i,j=1}^{l} \sum_{\alpha=l+1}^{n} \overline{\omega_{i\alpha}} \Big\langle [\overline{\mathfrak{e}}_1 \dots \overline{\mathfrak{e}}_{i-1}\, \overline{\mathfrak{e}}_\alpha\, \overline{\mathfrak{e}}_{i+1} \dots \overline{\mathfrak{e}}_l] \times \\
&\quad \times \Big(\omega_{jj}\, [\mathfrak{e}_1 \dots \mathfrak{e}_l] + \sum_{\beta=l+1}^{n} \omega_{j\beta}\, [\mathfrak{e}_1 \dots \mathfrak{e}_{j-1}\, \mathfrak{e}_\beta\, \mathfrak{e}_{j+1} \dots \mathfrak{e}_l]\Big)\Big\rangle \\
&= \lambda \sum_{i,j=1}^{l} \sum_{\alpha,\beta=l+1}^{n} \overline{\omega_{i\alpha}}\, \omega_{j\beta} \langle [\overline{\mathfrak{e}}_1 \dots \overline{\mathfrak{e}}_{i-1}\, \overline{\mathfrak{e}}_\alpha\, \overline{\mathfrak{e}}_{i+1} \dots \overline{\mathfrak{e}}_l]\, [\mathfrak{e}_1 \dots \mathfrak{e}_{j-1}\, \mathfrak{e}_\beta\, \mathfrak{e}_{j+1} \dots \mathfrak{e}_l]\rangle \\
&= \lambda \sum_{i=1}^{l} \sum_{\alpha=l+1}^{n} \overline{\omega_{i\alpha}}\, \omega_{i\alpha} = \lambda\Big(\sum_{i=1}^{l}\sum_{\alpha=l+1}^{n} |\langle \overline{\mathfrak{e}}_\alpha\, d\mathfrak{e}_i\rangle|^2\Big) = ds^2. \qquad \text{q.e.d.}
\end{aligned}$$

4.5. Gelegentlich werden in der Differentialgeometrie auch nicht notwendig positiv definite, nichtausgeartete quadratische Differentialformen[49]) in einer differenzierbaren Mannigfaltigkeit betrachtet. So liegt es nahe, in unserem Falle nichtausgeartete, gegen $SO(n)$ sowie gegen $O(n)$ bzw. gegen $SU(n)$ sowie gegen $U(n)$ invariante quadratische bzw. Hermitesche Differentialformen in $G_{l,m}$ und $\widetilde{G_{l,m}}$ bzw. $H_{l,m}$ zu untersuchen. Dabei zeigt es sich, daß die in Satz 1, Satz 2, Satz 3 und Satz 6 angegebenen Differentialformen auch die einzig möglichen *nichtausgearteten* quadratischen bzw. Hermiteschen Differentialformen mit den geforderten Eigenschaften sind, wenn man in diesen Sätzen nur die Konstanten λ und μ an Stelle von $\lambda>0$ und $\lambda>|\mu|$ den schwächeren Bedingungen $\lambda\neq 0$ und $|\lambda|\neq|\mu|$ unterwirft. (Zum Beweis dieser Tatsache braucht lediglich Hilfssatz 1 geeignet modifiziert zu werden!).

4.6. Neben den in 1. und 4.2 definierten Grassmannschen Mannigfaltigkeiten $G_{l,m}$, $\widetilde{G_{l,m}}$ und $H_{l,m}$ spielt (insbesondere in der Integralgeometrie) auch die Menge $I_{l,m}$ aller l-dimensionalen *linearen Unterräume* A_l des n-dimensionalen reellen *affinen Raums* A_n eine Rolle ($m=n-l$; $l\geqq 1$, $m\geqq 1$). Ein „Punkt" o von $I_{l,m}$ wird etwa in Parameterdarstellung (bei einem geeigneten Koordinatensystem von A_n) durch

$$\mathfrak{x}=\sum_{i=1}^{l}\xi_i\,\mathring{e}_i \tag{43}$$

dargestellt, wobei mit $\mathring{e}_i$ $(i=1,\ldots,l)$ l linear unabhängige Vektoren des n-dimensionalen Vektorraums V_n bezeichnet sind. Ergänzt man dieses l-Bein mittels der Vektoren $\mathring{e}_\alpha$ $(\alpha=l+1,\ldots,n)$ zu einem n-Bein von V_n, so kann jeder „Nachbarpunkt" von o in Parameterdarstellung in der Form

$$\mathfrak{x}=\Big(\sum_{\alpha=l+1}^{n}u_\alpha\,\mathring{e}_\alpha\Big)+\sum_{i=1}^{l}\xi_i\Big(\mathring{e}_i+\sum_{\alpha=l+1}^{n}u_{i\alpha}\,\mathring{e}_\alpha\Big) \tag{44}$$

(vgl. (1)!) angegeben werden. Hierbei gehören zu verschiedenen Koeffizientengesamtheiten $(u_\alpha, u_{i\alpha})$ verschiedene lineare Unterräume von A_n, d.h. verschiedene Punkte von $I_{l,m}$, so daß sich $I_{l,m}$ als eine reell analytische Mannigfaltigkeit der Dimension $(l+1)m$ mit den (durch das n-Bein $\{\mathring{e}_i, \mathring{e}_\alpha\}$ und den Nullpunkt des A_n bestimmten) lokalen Koordinaten u_α und $u_{i\alpha}$ $(i=1,\ldots,l;$ $\alpha=l+1,\ldots,n)$ erweist.

In Analogie zu den früher betrachteten Grassmannschen Mannigfaltigkeiten wollen wir nun auf $I_{l,m}$ alle Riemannschen Metriken zu bestimmen suchen, welche gegenüber derjenigen Gruppe invariant sind, die von der Gruppe $SB(n)$ aller eigentlichen Bewegungen des (mit einer euklidischen Struktur versehenen) Raums A_n auf $I_{l,m}$ induziert wird. Zu diesem Zweck denken wir

[49]) Unter einer „nichtausgearteten" quadratischen Differentialform bzw. Hermiteschen Differentialform versteht man eine quadratische bzw. Hermitesche Differentialform, bei welcher die aus ihren Koeffizienten gebildete Determinante von Null verschieden ist.

uns das n-Bein $\{\mathring{e}_i, \mathring{e}_\alpha\}$ orthonormiert. Dann führt ein beliebiges, den linearen Unterraum (43) als Ganzes festlassendes Element von $SB(n)$ den benachbarten linearen Unterraum (44) in den linearen Unterraum

$$(45)\quad \mathfrak{x} = \sum_{i=1}^{l} a_i \mathring{e}_i + \sum_{\alpha=l+1}^{n} u_\alpha \Big(\sum_{\beta=l+1}^{n} b_{\beta\alpha} \mathring{e}_\beta \Big) + \sum_{i=1}^{l} \xi_i \Big(\sum_{j=1}^{l} a_{ji} \mathring{e}_j + \sum_{\alpha=l+1}^{n} u_{i\alpha} \Big(\sum_{\beta=l+1}^{n} b_{\beta\alpha} \mathring{e}_\beta \Big) \Big)$$

über, wobei (a_{ij}) und $(b_{\alpha\beta})$ zwei orthogonale Matrizen sind, deren Determinanten $\|a_{ij}\|$ und $\|b_{\alpha\beta}\|$ gleiches Vorzeichen besitzen, während die l Größen a_i keiner Einschränkung unterliegen. Der lineare Unterraum (45) gestattet, wie eine leichte Umrechnung zeigt, auch die Parameterdarstellung:

$$\mathfrak{x} = \Big(\sum_{\alpha=l+1}^{n} u'_\alpha \mathring{e}_\alpha \Big) + \sum_{i=1}^{l} \eta_i \Big(\mathring{e}_i + \sum_{\alpha=l+1}^{n} u'_{i\alpha} \mathring{e}_\alpha \Big)$$

mit

$$u'_\alpha = \sum_{\beta=l+1}^{n} b_{\alpha\beta} u_\beta - \sum_{j,k=1}^{l} \sum_{\beta=l+1}^{n} a_k a_{kj} b_{\alpha\beta} u_{j\beta},$$

$$u'_{i\alpha} = \sum_{j=1}^{l} \sum_{\beta=l+1}^{n} a_{ij} b_{\alpha\beta} u_{j\beta} \qquad (i = 1, \ldots, l;\ \alpha = l+1, \ldots, n).$$

Daraus folgt, daß die (in dem Tangentialraum U von $I_{l,m}$ im Punkt o operierende) lineare Isotropiegruppe Δ von $I_{l,m}$ bezüglich $SB(n)$ aus den linearen Transformationen

$$(46)\quad \begin{cases} x'_\alpha = \sum_{\beta=l+1}^{n} b_{\alpha\beta} x_\beta - \sum_{j,k=1}^{l} \sum_{\beta=l+1}^{n} a_k a_{kj} b_{\alpha\beta} x_{j\beta} \\ x'_{i\alpha} = \sum_{j=1}^{l} \sum_{\beta=l+1}^{n} a_{ij} b_{\alpha\beta} x_{j\beta} \qquad (i = 1, \ldots, l;\ \alpha = l+1, \ldots, n) \end{cases}$$

besteht (vgl. (4); x_α und $x_{i\alpha}$ = Komponenten der Vektoren von U!).

Die Bestimmung aller gegenüber $SB(n)$ invarianten Riemannschen Metriken von $I_{l,m}$ erweist sich jetzt wieder als äquivalent mit der Angabe aller positiv definiten und gegenüber Δ invarianten quadratischen Formen der Variablen x_α und $x_{i\alpha}$. Nun existiert allerdings keine einzige derartige quadratische Form, da anderenfalls die von den Translationen das A_n in $I_{l,m}$ induzierte und aus den Elementen

$$(47)\quad \begin{cases} x'_\alpha = x_\alpha - \sum_{k=1}^{l} a_k x_{k\alpha} \\ x'_{i\alpha} = x_{i\alpha} \qquad (i = 1, \ldots, l;\ \alpha = l+1, \ldots, n) \end{cases}$$

bestehende Untergruppe Δ_1 von Δ als abgeschlossene Untergruppe der kompakten Gruppe aller die quadratische Form invariant lassenden linearen Transformationen ebenfalls kompakt sein müßte, was offensichtlich nicht zutrifft.

Infolgedessen gibt es überhaupt keine gegen $SB(n)$ invariante Riemannsche Metrik von $I_{l,m}$. Damit ist allerdings noch nichts über die Existenz von gegen $SB(n)$ invarianten, nichtausgearteten quadratischen Differentialformen in $I_{l,m}$ gesagt, worüber der folgende Satz Auskunft gibt:

SATZ 8. *Nichtausgeartete und gegen $SB(n)$ invariante quadratische Differentialformen in $I_{l,m}$ sind nur im Fall $l=1$ und $m=2$ vorhanden; ihre Gesamtheit wird dargestellt durch:*

$$2\lambda(\langle \mathfrak{e}_2\, d\mathfrak{p}\rangle\langle \mathfrak{e}_3\, d\mathfrak{e}_1\rangle - \langle \mathfrak{e}_3\, d\mathfrak{p}\rangle\langle \mathfrak{e}_2\, d\mathfrak{e}_1\rangle) + \mu\left((\langle \mathfrak{e}_2\, d\mathfrak{e}_1\rangle)^2 + (\langle \mathfrak{e}_3\, d\mathfrak{e}_1\rangle)^2\right)^{50})$$
$$(\lambda, \mu = \text{reelle Konstante};\ \lambda \neq 0).$$

BEWEIS. Nach einer vorigen Bemerkung und auf Grund der Bedeutung von $\mathfrak{p}, \mathfrak{e}_1, \mathfrak{e}_2, \mathfrak{e}_3$ genügt es nachzuweisen, daß im Fall $l=1$, $m=2$

$$2\lambda(x_2 x_{13} - x_3 x_{12}) + \mu\left((x_{12})^2 + (x_{13})^2\right) \qquad (\lambda \neq 0) \tag{48}$$

genau die gegen Δ invarianten, nichtausgearteten quadratischen Formen auf dem Tangentialraum von $I_{l,m}$ in o sind, während für andere Wertepaare von l und m überhaupt keine derartigen quadratischen Formen existieren. Es ist dabei zweckmäßig, zunächst den Nachweis für die letztere Behauptung zu erbringen.

Dazu nehmen wir einmal an,

$$\left\{\sum_{\alpha,\beta=l+1}^{n} c_{(\alpha)(\beta)}\, x_\alpha x_\beta + 2\sum_{i=1}^{l}\sum_{\alpha,\beta=l+1}^{n} c_{(\alpha)(i\beta)}\, x_\alpha x_{i\beta} + \sum_{i,j=1}^{l}\sum_{\alpha,\beta=l+1}^{n} c_{(i\alpha)(j\beta)}\, x_{i\alpha} x_{j\beta}\right. \tag{49}$$
$$\left.(c_{(\alpha)(\beta)} = c_{(\beta)(\alpha)},\ c_{(\alpha)(i\beta)} = c_{(i\beta)(\alpha)},\ c_{(i\alpha)(j\beta)} = c_{(j\beta)(i\alpha)})\right.$$

sei irgendeine nichtausgeartete und gegenüber allen Transformationen (46) invariante quadratische Form der x_α und $x_{i\alpha}$. Dann ist diese quadratische Form auch um so mehr gegen die Transformationen (47) invariant. Daraus folgt aber durch Einsetzen mit nachfolgendem Koeffizientenvergleich:

$$c_{(\alpha)(\beta)} = 0 \quad \text{und} \quad c_{(\alpha)(i\beta)}\, a_j + c_{(\beta)(j\alpha)}\, a_i = 0.$$
$$(i, j = 1, \ldots, l;\ \alpha, \beta = l+1, \ldots, n).$$

Wäre nun $l>1$, so ergibt die zweite Gleichung weiter $c_{(\alpha)(i\beta)}=0$ $(i=1, \ldots, l;$ $\alpha, \beta = l+1, \ldots, n)$ (die a_i sind beliebig wählbar!), so daß sich (49) auf $\sum_{i,j=1}^{l}\sum_{\alpha,\beta=l+1}^{n} c_{(i\alpha)(j\beta)} x_{i\alpha} x_{j\beta}$ reduziert. Dies steht aber mit der Voraussetzung im Widerspruch, daß unsere quadratische Form nichtausgeartet sein sollte. Infolgedessen ist notwendig $l=1$ sowie $m>1$, und (49) besitzt die Gestalt:

$$2\sum_{2 \leq \alpha < \beta \leq n} c_{(\alpha)(1\beta)}(x_\alpha x_{1\beta} - x_\beta x_{1\alpha}) + \sum_{\alpha,\beta=2}^{n} c_{(1\alpha)(1\beta)}\, x_{1\alpha} x_{1\beta}. \tag{50}$$

[50]) Hierbei haben die Vektorfunktionen $\mathfrak{e}_1, \mathfrak{e}_2, \mathfrak{e}_3$ die analoge Bedeutung wie auf S. 347, und die Vektorfunktion $\mathfrak{p}$ stellt den Ortsvektor eines beliebigen Punktes auf den Geraden von $I_{1,2}$ dar. Satz 8 enthält übrigens die schon bewiesene Tatsache, daß es keine gegen $SB(n)$ invariante Riemannsche Metrik in $I_{l,m}$ gibt.

Wäre jetzt $m \geqq 3$, so müßte wegen der Invarianz von (50) gegen die durch geeignete Spezialisierung der Koeffizienten in (46) erhaltene Transformation $x'_{\alpha_0} = x_{\beta_0}$, $x'_{1\alpha_0} = x_{1\beta_0}$, $x'_{\beta_0} = x_{\alpha_0}$, $x'_{1\beta_0} = x_{1\alpha_0}$, $x'_{\gamma_0} = -x_{\gamma_0}$, $x'_{1\gamma_0} = -x_{1\gamma_0}$, $x'_\delta = x_\delta$, $x'_{1\delta} = x_{1\delta}$ ($\alpha_0 < \beta_0$, $\gamma_0 \neq \alpha_0$, $\gamma_0 \neq \beta_0$, δ durchläuft alle von den speziellen Indizes $\alpha_0, \beta_0, \gamma_0$ verschiedenen Indizes zwischen 2 und n) $c_{(\alpha_0)(1\beta_0)} = 0$ gelten. Da aber hierbei das Indexpaar α_0, β_0 bis auf die Einschränkung $2 \leqq \alpha_0 < \beta_0 \leqq n$ beliebig gewählt werden kann, reduziert sich (50) auf $\sum_{\alpha,\beta=2}^{n} c_{(1\alpha)(1\beta)}\, x_{1\alpha}\, x_{1\beta}$, was wieder wie vorhin unmöglich ist. Also gilt in der Tat $l = 1$ sowie $m = 2$, und unsere quadratische Form lautet:

$$2\, c_{(2)(13)}(x_2\, x_{13} - x_3\, x_{12}) + \sum_{\alpha,\beta=2}^{3} c_{(1\alpha)(1\beta)}\, x_{1\alpha}\, x_{1\beta}. \tag{51}$$

Jetzt lehrt einfaches Nachrechnen, daß sich der erste Summand von (51) bei allen Transformationen (46) nicht ändert. Da dies aber auch für die ganze quadratische Form (51) zutreffen muß, folgt dasselbe für den zweiten Summanden von (51), der infolgedessen auf Grund der in Abschnitt 2 durchgeführten Betrachtungen die Gestalt $\mu\left((x_{12})^2 + (x_{13})^2\right)$ besitzt. Berücksichtigt man nun noch, daß wegen des Nichtausgeartetseins von (51) $c_{(2)(13)} \neq 0$ gilt, so findet man schließlich für unsere quadratische Form den (tatsächlich eine nichtausgeartete, gegen Δ invariante quadratische Form darstellenden) Ausdruck (48), womit der Beweis von Satz 8 beendet ist. q.e.d.

Der Beweis von Satz 8 liefert nach leichter Abänderung zusätzlich die Tatsache, *daß es keine gegen die Gruppe $B(n)$ aller eigentlichen und uneigentlichen Bewegungen des A_n invariante, nichtausgeartete quadratische Differentialform in $I_{l,m}$ geben kann.* Außerdem soll noch erwähnt werden, daß die Koeffizientendeterminanten aller Elemente (46) der linearen Isotropiegruppe von $I_{l,m}$ bezüglich $SB(n)$ vom Betrag 1 sind; dasselbe gilt auch für die lineare Isotropiegruppe von $I_{l,m}$ bezüglich $B(n)$. *Aus diesem Grunde existieren gegen $B(n)$ invariante Volumelemente von $I_{l,m}$, welche sich alle in der Form:*

$$dV = \lambda\left(\left(\bigwedge_{\alpha=l+1}^{n} \langle \mathfrak{e}_\alpha\, d\mathfrak{p}\rangle\right) \wedge \left(\bigwedge_{i=1}^{l} \bigwedge_{\alpha=l+1}^{n} \langle \mathfrak{e}_\alpha\, d\mathfrak{e}_i\rangle\right)\right) \qquad (\lambda = \text{const} \neq 0)$$

angeben lassen[51]).

Literatur

[1] Whitney, H.: Differentiable manifolds. Ann. Math. **37**, 645–680 (1936).

[2] Teleman, C.: Sur les variétés de Grassmann. Bull. math. Soc. sci. math. phys. Rép. Popul. Roumaine 2 (50), 203–224 (1958).

[3] Bourbaki, N.: Algèbre ch. III. Paris 1948.

[4] Boerner, H.: Darstellungen von Gruppen. Berlin 1955.

[5] Chern, S. S.: Topics in differential geometry (mimeogr. notes). Princeton: Inst. of Adv. Study 1951.

[51]) Vgl. Fußnote [50]). Diese Volumelemente spielen in der Integralgeometrie eine Rolle; man findet sie im Fall $l = 1$, $m = 2$ und $l = 2$, $m = 1$ bei [16], II (33) und (24).

[6] EISENHART, L. P.: Riemannian geometry. Princeton 1949.
[7] CARTAN, É.: Oeuvres complètes. I. Paris 1952.
[8] NOMIZU, K.: Invariant affine connections on homogeneous spaces. Amer. J. Math. **76**, 33–65 (1954).
[9] HELGASON, S.: On Riemannian curvature of homogeneous spaces. Proc. Amer. Math. Soc. **9**, 831–838 (1958).
[10] PONTRJAGIN, L.: Topological groups. Princeton 1946.
[11] WEYL, H.: Gruppentheorie und Quantenmechanik. Leipzig 1928.
[12] EHRESMANN, C.: Sur la topologie de certains espaces homogènes. Ann. Math. **35**, 396–443 (1934).
[13] KÄHLER, E.: Über eine bemerkenswerte Hermitesche Metrik. Hamb. Abh. **9**, 173–186 (1933).
[14] CHERN, S. S.: Characteristic classes of Hermitian manifolds. Ann. Math. **47**, 85–121 (1946).
[15] HODGE, W. U. P., u. D. PEDOE: Methods of algebraic geometry. I. Cambridge 1953.
[16] BLASCHKE, W.: Vorlesungen über Integralgeometrie. II. Leipzig 1937.

Freiburg i. Br., Mathematisches Institut der Universität

(Eingegangen am 26. Oktober 1960)

BEHRENS, E.-A.
Math. Zeitschr. 76, 367–384 (1961)

Distributiv darstellbare Ringe. II

Von
ERNST-AUGUST BEHRENS

Als distributiv darstellbar wurde ein Artin-Ring $\mathfrak{R}$ mit Einselement in BEHRENS [4][1]) definiert, wenn es einen treuen $\mathfrak{R}$-Modul $\mathfrak{m}$ mit distributivem $\mathfrak{R}$-Untermodulverband $V(\mathfrak{m})$ gibt. Die dort in erster Linie für primäre Ringe begonnenen Untersuchungen sollen in dieser Abhandlung fortgesetzt werden. Bereits in Teil I erwies es sich als zweckmäßig, den Begriff der distributiven Darstellbarkeit primärer Ringe $\mathfrak{R}=K_d+\mathfrak{W}$ zu dem der distributiven und K-linearen Darstellbarkeit zu verschärfen, indem man verlangt, daß $\mathfrak{m}$ ein $\mathfrak{R}$-Links-, K-Rechtsmodul ist. Dazu werden wir in §1, Satz 1 sehen, daß jeder $\mathfrak{R}$-Untermodul eines $\mathfrak{R}$-K-Moduls $\mathfrak{m}$ ein K-Rechtsmodul ist, falls der $\mathfrak{R}$-Untermodulverband $V(\mathfrak{m})$ distributiv ist. Dies gilt sogar für nicht notwendig mehr primäre, distributiv darstellbare Artin-Ringe $\mathfrak{R}$ mit Einselement, wenn K ein im Zentralisator des $\mathfrak{R}$-Moduls $\mathfrak{m}$ enthaltener Ring ist. Im folgenden sei unter einem distributiv darstellbaren Ring $\mathfrak{R}$ stets ein distributiv und für einen geeigneten Ring K auch K-linear darstellbarer Ring verstanden.

Ein distributiver $\mathfrak{R}$-K-Modul $\mathfrak{m}$ ist zwar im allgemeinen nicht eindeutig bis auf $\mathfrak{R}$-Modulisomorphie durch die Ringe $\mathfrak{R}$ und K bestimmt, wohl aber sein Verband $V(\mathfrak{m})$ bis auf Verbandsisomorphie. Diese für primäre Ringe, bei denen ja $V(\mathfrak{m})$ eine Kette ist, bereits in Teil I enthaltene Erkenntnis wird in §1 aus einer Beschreibung des Verbandes $V(\mathfrak{m})$ innerhalb des Ringes $\mathfrak{R}$ abgeleitet werden.

Ein primärer distributiv darstellbarer Ring $\mathfrak{R}$ besitzt nach Teil I, Satz 3, genau ein minimales Ideal, nämlich $\mathfrak{W}^{e-1}$, wobei e der Exponent des Radikales $\mathfrak{W}$ von $\mathfrak{R}$ ist. Auch ein beliebiger distributiv darstellbarer Ring $\mathfrak{R}$ enthält nur wenige minimale Ideale, die in §2, Satz 5, innerhalb des Ringes $\mathfrak{R}$ beschrieben werden. Jedoch erläutert erst dessen Zusatz, daß diese minimalen Ideale eng mit der Struktur der von den verbindungsirreduziblen $\mathfrak{R}$-Untermoduln eines distributiven $\mathfrak{R}$-Moduls $\mathfrak{m}$ gebildeten Halbordnung zusammenhängen.

Jeder distributive $\mathfrak{R}$-K-Modul $\mathfrak{m}$ ist zyklisch, $\mathfrak{m}=\mathfrak{R}\,b$, und damit isomorph zu $\mathfrak{R}-(0:b)$, wobei das Linksideal $\mathfrak{l}=(0:b)$ einige, bereits in $\mathfrak{R}$ ausdrückbare charakteristische Eigenschaften besitzt. Die distributiv darstellbaren Ringe $\mathfrak{R}$ sind dann natürlich diejenigen, die ein Linksideal $\mathfrak{l}$ mit diesen Eigenschaften enthalten (Satz 6). Diese im Grunde selbstverständliche ringinterne Charakterisierung der distributiv darstellbaren Ringe wird wesentlich interessanter, wenn man $\mathfrak{l}$ innerhalb des Ringes $\mathfrak{R}$ und seiner Faktorringe

[1]) Im folgenden wird diese Arbeit als „Teil I" zitiert.

durch Maximaleigenschaften näher beschreibt. In §3, Satz 7, wird diese Aufgabe für beliebige Ringe auf das analoge Problem in primären Ringen zurückgeführt. Für primäre Ringe sind nun die zu distributiven $\mathfrak{R}$-K-Moduln $\mathfrak{R}-\mathfrak{l}$ führenden Linksideale $\mathfrak{l}$ bereits aus Teil I, §2, näher bekannt. Damals mußten aber über die Linksideale in $\mathfrak{R}$ hinaus noch die ω-Untermoduln von $\mathfrak{R}$ herangezogen werden, wobei ω ein Element aus $\mathfrak{W}\setminus\mathfrak{W}^2$ mit bestimmten Eigenschaften ist. Die Existenz dieses ω ist zwar in jedem distributiv darstellbaren primären Ring gesichert, will man aber diese Ringe durch innere Eigenschaften charakterisieren, muß man die Existenz von ω in $\mathfrak{R}$ postulieren, wie in Teil I, am Ende von §2, näher ausgeführt worden ist. In §4, Satz 8, werden wir nun eine Beschreibung der zu distributiven $\mathfrak{R}$-K-Moduln $\mathfrak{R}-\mathfrak{l}$ führenden Linksideale kennenlernen, in welcher nur Maximaleigenschaften von $\mathfrak{l}$ als Linksideal aber nicht mehr als ω-Modul auftreten. Dies erlaubt, zusammen mit Satz 7, eine durchsichtigere ringinterne Charakterisierung der distributiv darstellbaren Ringe anzugeben.

§1

$\mathfrak{R}$ sei im folgenden stets ein Artin-Ring mit Einselement, d.h. in $\mathfrak{R}$ sei die Minimalbedingung für Linksideale erfüllt. $\mathfrak{W}$ sei sein Radikal und e dessen Exponent. Die Struktur eines endlichen distributiven Verbandes V läßt sich bekanntlich (HERMES [6], §21, BIRKHOFF [5], X) erfassen durch die Halbordnungsrelation zwischen seinen verbindungsirreduziblen Elementen. Dabei heißt ein Verbandselement $\mathfrak{i}$ verbindungsirreduzibel, wenn es nicht die Verbindung $\mathfrak{a}_1\cup\mathfrak{a}_2$ zweier Verbandselemente $\mathfrak{a}_i$ mit $\mathfrak{a}_i<\mathfrak{i}$ ist. Diese Beschreibung des Verbandes V ist möglich, weil jedes seiner Elemente Verbindung von verbindungsirreduziblen ist und weiter ein verbindungsirreduzibles Element $\mathfrak{i}$ genau dann in einer Verbindung der Elemente $\mathfrak{a}_i$ aus V enthalten ist, wenn $\mathfrak{i}$ in mindestens einem der $\mathfrak{a}_i$ liegt. In unserem Falle ist V der Verband der $\mathfrak{R}$-Untermoduln eines $\mathfrak{R}$-Moduls $\mathfrak{m}$ mit der Inklusion als Halbordnungsrelation.

$\mathfrak{R}$ ist nach der Definition in Teil I distributiv darstellbar, wenn $\mathfrak{R}$ einen treuen $\mathfrak{R}$-Modul $\mathfrak{m}$ mit distributivem $\mathfrak{R}$-Untermodulverband $V(\mathfrak{m})$ besitzt. Für primäre Ringe erwies es sich als nützlich, den Begriff der distributiven und K-linearen Darstellbarkeit von $\mathfrak{R}$ durch die Forderung der Existenz eines treuen $\mathfrak{R}$-Links-, K-Rechtsmoduls $\mathfrak{m}$ mit distributivem $\mathfrak{R}$-Untermodulverband $V(\mathfrak{m})$ einzuführen. Der primäre Ring $\mathfrak{R}$ und der Körper K standen dabei in der Beziehung $\mathfrak{R}/\mathfrak{W}=K_d$ zueinander. Etwas allgemeiner gelte folgende

DEFINITION 1. *Die Ringe $\mathfrak{R}$ und K seien Artin-Ringe mit Einselementen. $\mathfrak{R}$ ist distributiv und K-linear darstellbar, wenn ein als $\mathfrak{R}$-Modul treuer und für $\mathfrak{R}$ und K unitaler $\mathfrak{R}$-Links-, K-Rechtsmodul $\mathfrak{m}$ mit distributivem $\mathfrak{R}$-Untermodulverband $V(\mathfrak{m})$ existiert. $\mathfrak{m}$ sei in diesem Falle als distributiver $\mathfrak{R}$-K-Modul bezeichnet.*

Der folgende Satz 1 besagt, daß man in dieser Definition statt des $\mathfrak{R}$-Untermodulverbandes auch den $\mathfrak{R}$-K-Untermodulverband von $\mathfrak{m}$ hätte nehmen können.

SATZ 1. *Jeder $\mathfrak{R}$-Untermodul eines distributiven $\mathfrak{R}$-K-Moduls ist ebenfalls ein K-Rechtsmodul.*

Nimmt man daher insbesondere für K den ganzen $\mathfrak{R}$-Zentralisator $\mathfrak{Z}$ des distributiven $\mathfrak{R}$-Moduls $\mathfrak{m}$, so ist jeder $\mathfrak{R}$-Untermodul von $\mathfrak{m}$ vollinvariant im Sinne der Definition von JACOBSON [7], IV. 2.2.

BEWEIS DES SATZES 1. Angenommen, der $\mathfrak{R}$-Untermodul $\mathfrak{u}$ wäre kein K-Rechtsmodul, also $\mathfrak{u}K > \mathfrak{u}$. Da $\mathfrak{W}^0\mathfrak{u} = \mathfrak{u}$, $\mathfrak{W}^e\mathfrak{u} = 0$, existiert eine Zahl p zwischen 0 und $e-1$ mit $\mathfrak{W}^p\mathfrak{u}K > \mathfrak{W}^p\mathfrak{u}$, aber $\mathfrak{W}^{p+1}\mathfrak{u}K = \mathfrak{W}^{p+1}\mathfrak{u}$. Mit $\mathfrak{v} = \mathfrak{W}^{p+1}\mathfrak{u}$ ist auch $\overline{\mathfrak{m}} = \mathfrak{m} - \mathfrak{v}$ ein K-Rechtsmodul also ist $\mathfrak{W}^p\overline{\mathfrak{u}}K = \mathfrak{W}^p\mathfrak{u}K - \mathfrak{v}$ ein $\mathfrak{R}/\mathfrak{W}$-$K$-Modul. Die Distributivität der $\mathfrak{R}$-Untermodulverbände des Faktormoduls $\overline{\mathfrak{m}}$ und seines Untermoduls $\mathfrak{W}^p\overline{\mathfrak{u}}K$ folgt aus der von $V(\mathfrak{m})$. Auch für die Faktormoduln gilt $\mathfrak{W}^p\overline{\mathfrak{u}}K > \mathfrak{W}^p\overline{\mathfrak{u}}$, da sonst $\mathfrak{W}^p\overline{\mathfrak{u}}K = \mathfrak{W}^p\mathfrak{u} \cup \mathfrak{v} = \mathfrak{W}^p\mathfrak{u} \cup \mathfrak{W}^{p+1}\mathfrak{u} = \mathfrak{W}^p\mathfrak{u}$ wäre. Also existiert ein Element $\varkappa \in K$ mit $\mathfrak{W}^p\overline{\mathfrak{u}}\varkappa \subsetneqq \mathfrak{W}^p\overline{\mathfrak{u}}$. Setze $\overline{\mathfrak{b}} = \mathfrak{W}^p\overline{\mathfrak{u}}\varkappa \cap \mathfrak{W}^p\overline{\mathfrak{u}}$. Als $\mathfrak{R}/\mathfrak{W}$-Modul ist $\mathfrak{W}^p\overline{\mathfrak{u}}$ vollreduzibel, also eine $\mathfrak{R}$-direkte Summe $\overline{\mathfrak{b}} + \overline{\mathfrak{b}}_1 + \cdots + \overline{\mathfrak{b}}_s$ mit irreduziblen $\mathfrak{R}$-Moduln $\overline{\mathfrak{b}}_i$. Für mindestens ein $\overline{\mathfrak{b}}_i$ ist dabei $\overline{\mathfrak{b}}_i\varkappa \subsetneqq \overline{\mathfrak{b}}_i$. Die Irreduzibilität von $\overline{\mathfrak{b}}_j$ impliziert, daß der $\mathfrak{R}$-Homomorphismus $H\colon \overline{\mathfrak{b}}_j \to \overline{\mathfrak{b}}_j\varkappa$ ein Isomorphismus ist. Demnach enthält $\mathfrak{W}^p\overline{\mathfrak{u}}\,K$ zwei verschiedene, zueinander isomorphe $\mathfrak{R}$-Moduln $\overline{\mathfrak{b}}_j$ und $\overline{\mathfrak{b}}_j\varkappa$, deren Durchschnitt der Nullmodul $\overline{0}$ ist. Dies widerspricht aber der Distributivität des Verbandes der $\mathfrak{R}$-Untermoduln von $\mathfrak{W}^p\overline{\mathfrak{u}}K$, da dieser dann auch den $\mathfrak{R}$-Modul $\overline{\mathfrak{c}} = \{\overline{b}_j + \overline{b}_j H \mid \overline{b}_j \in \overline{\mathfrak{b}}_j\}$, sein Diagramm also die folgende Figur enthielte:

In jedem Artin-Ring $\mathfrak{R}$ mit 1 existieren bekanntlich orthogonale Idempotente $\varepsilon_1, \ldots, \varepsilon_q$ mit $\varepsilon_1 + \cdots + \varepsilon_q = 1$ derart, daß der Ring als additive Gruppe die direkte Summe

$$\mathfrak{R} = \mathfrak{R}_1 + \cdots + \mathfrak{R}_q + \mathfrak{N} \tag{1}$$

ist, deren Summanden $\mathfrak{R}_s = \varepsilon_s\mathfrak{R}\varepsilon_s$ primäre Ringe sind und in der $\mathfrak{N}$ eine additive Untergruppe des Radikales $\mathfrak{W}$ des Ringes $\mathfrak{R}$ ist. Modulo $\mathfrak{W}$ betrachtet führt (1) auf die direkte Summendarstellung des halbeinfachen Ringes $\mathfrak{R}/\mathfrak{W}$ durch einfache Ringe $\overline{\mathfrak{R}}_s$ im ringtheoretischen Sinne. $\mathfrak{W}_s$ sei das Radikal von $\mathfrak{R}_s$ und e_s sein Exponent. Also existiert zu jedem s ein Körper $K^{(s)}$ mit

$$\overline{\mathfrak{R}}_s = \mathfrak{R}_s/\mathfrak{W}_s \cong K^{(s)}_{d_s}. \tag{2}$$

Falls $\mathfrak{R}$ primär, also $q = 1$ ist, besteht der Verband $V(\mathfrak{m})$ nach Teil I, §1, aus der Kette der $\mathfrak{W}^a_1\mathfrak{m}$, $a = 0, 1, \ldots, e-1$, so daß die verbandstheoretische Struktur von $V(\mathfrak{m})$ sich durch die Inklusionen in der Kette der Ideale $\mathfrak{W}^a_1$ erfassen läßt. Das nächste Ziel ist, auch für $q > 1$ eine Beschreibung von

$V(\mathfrak{m})$ innerhalb des Ringes $\mathfrak{R}$ anzugeben. Dabei legen es die Überlegungen in BEHRENS [2] nahe, folgendermaßen vorzugehen:

HILFSSATZ 1. *Sei*

$$0 = \mathfrak{m}_0 < \mathfrak{m}_1 < \cdots < \mathfrak{m}_l = \mathfrak{m} \tag{3}$$

eine maximale Kette von $\mathfrak{R}$-Untermoduln des distributiven $\mathfrak{R}-K$-Moduls $\mathfrak{m}$. Jeder der $\mathfrak{R}$-irreduziblen Faktormoduln $\mathfrak{m}_i - \mathfrak{m}_{i-1}$ stellt demnach genau einen der einfachen Ringe $\mathfrak{R}_s$ aus (2) treu dar, etwa $\mathfrak{R}_r$, und annulliert die anderen. Setze $r = f(i)$. Zu jedem $i = 1, \ldots, l$ existiert eine additive Untergruppe $\mathfrak{d}_i$ von $\mathfrak{m}$ mit $\varepsilon_r \mathfrak{d}_i = \mathfrak{d}_i$ derart, daß $\mathfrak{m}_i$ als Gruppe die direkte Summe von $\mathfrak{m}_{i-1}$ und $\mathfrak{d}_i$ ist. Setze $c_{ik} = 0$, wenn $\mathfrak{d}_i \cap \mathfrak{R}\mathfrak{d}_k = 0$ ist, und andernfalls $c_{ik} = 1$:

Behauptung.

$$\mathfrak{i}_k = \sum_{c_{ik}=1} \mathfrak{d}_i = \mathfrak{R}\mathfrak{d}_k, \qquad k = 1, \ldots, l, \tag{4}$$

sind die verbindungsirreduziblen $\mathfrak{R}$-Untermoduln von $\mathfrak{m}$. Bezüglich ihrer Inklusionen gilt:

$$\mathfrak{i}_j \leqq \mathfrak{i}_k \quad \textit{ist äquivalent} \quad c_{jk} = 1. \tag{5}$$

BEWEIS. Die Existenz der obigen $\mathfrak{d}_i$ in $\mathfrak{m} = \mathfrak{d}_1 + \cdots + \mathfrak{d}_l$ folgt daraus, daß $\mathfrak{m}_i - \mathfrak{m}_{i-1}$ ein $K^{(r)}$-Rechtsmodul ist: Seien $b_1, \ldots, b_d$ eine $K^{(r)}$-Basis, wobei alle $b_m \in \mathfrak{m}_i$. Dann enthält $\mathfrak{m}_i$ die als Gruppe zu $\mathfrak{m}_i - \mathfrak{m}_{i-1}$ isomorphe Gruppe $b_1 K^{(r)} + \cdots + b_d K^{(r)} = \mathfrak{d}_i'$. Also ist $\mathfrak{m}_i = \mathfrak{m}_{i-1} + \mathfrak{d}_i'$. Da ε_r den Modul $\mathfrak{m}_i - \mathfrak{m}_{i-1}$ elementweise reproduziert, ist $(1 - \varepsilon_r)\mathfrak{d}_i' \leqq \mathfrak{m}_{i-1}$ und demnach $\mathfrak{d}_i = \varepsilon_r \mathfrak{d}_i'$ als Gruppe direkter Summand in $\mathfrak{m}_i = \mathfrak{m}_{i-1} + \mathfrak{d}_i$ mit $\varepsilon_r \mathfrak{d}_i = \mathfrak{d}_i$. — Nach BEHRENS [3], S. 268–269, läßt sich jeder $\mathfrak{R}$-Untermodul $\mathfrak{u}$ von $\mathfrak{m}$ durch die als additive Gruppe direkte Summe

$$\mathfrak{u} = \sum_{i=1}^{q} \mathfrak{W}_i^{u_i} \mathfrak{m} \tag{6}$$

mit eindeutig bestimmten Exponenten u_i zwischen 0 und e_i ausdrücken. Analog seien die Zahlen v_i die Exponenten von $\mathfrak{v} \in V(\mathfrak{m})$. Dann gehören zu der Verbindung $\mathfrak{u} \cup \mathfrak{v}$ die Exponenten $s_i = \operatorname{Min}(u_i, v_i)$. Angenommen nun, $\mathfrak{i}_k = \mathfrak{R}\mathfrak{d}_k$ wäre reduzibel, also $\mathfrak{i}_k = \mathfrak{u} \cup \mathfrak{v}$, aber $\mathfrak{i}_k$ von $\mathfrak{u}$ und $\mathfrak{v}$ verschieden. Mit der Bezeichnung $s = f(k)$ ist $\mathfrak{i}_k = \mathfrak{R}\mathfrak{d}_k \leqq \mathfrak{u}$ äquivalent mit $\mathfrak{d}_k = \varepsilon_s \mathfrak{d}_k \leqq \mathfrak{u}$, also auch mit $\mathfrak{d}_k \leqq \varepsilon_s \mathfrak{u} = \mathfrak{W}_s^{u_s} \mathfrak{m}$, denn da $\mathfrak{R}_s$ ein primärer Ring ist, bilden die $\mathfrak{R}_s$-Untermoduln von $\mathfrak{m}$ die Kette der $\mathfrak{W}_s^b \mathfrak{m}$, $0 \leqq b \leqq e_s$. Analog für $\mathfrak{v}$ und für $\mathfrak{u} \cup \mathfrak{v}$. Aus $\mathfrak{i}_k = \mathfrak{u} \cup \mathfrak{v}$ folgt demnach $\mathfrak{d}_k \leqq \mathfrak{W}_s^m \mathfrak{m}$, wobei $m = \min(u_s, v_s)$ ist. Aus $\mathfrak{i}_k \not\leqq \mathfrak{u}$ und $\mathfrak{i}_k \not\leqq \mathfrak{v}$ folgt andererseits $\mathfrak{d}_k \not\leqq \mathfrak{W}_s^{u_s} \mathfrak{m}$ und $\mathfrak{d}_k \not\leqq \mathfrak{W}_s^{v_s} \mathfrak{m}$, q.e.a. — Dies impliziert $\mathfrak{m}_i = \mathfrak{i}_1 \cup \cdots \cup \mathfrak{i}_i$. Sei x_i ein nicht in $\mathfrak{m}_{i-1}$ enthaltenes Element aus $\mathfrak{m}_i \cap \mathfrak{R}\mathfrak{d}_k$. Da auch der $\mathfrak{R}$-Modul $\mathfrak{R}x_i$ in diesem Durchschnitt liegt, tritt in dessen Darstellung als Verbindung von $\cup$-irreduziblen Elementen $\mathfrak{i}_i$ auf. Aus $\mathfrak{R}x_i \leqq \mathfrak{R}\mathfrak{d}_k$ folgt dann $\mathfrak{d}_i \leqq \mathfrak{R}\mathfrak{d}_k$. (Insbesondere ist $c_{ik} = 1$ äquivalent $\mathfrak{d}_i \leqq \mathfrak{R}\mathfrak{d}_k$.) Demnach enthält schließlich $\mathfrak{R}\mathfrak{d}_k$ mit $m_{i-1} + d_i$ auch die Elemente $m_{i-1} \in \mathfrak{m}_{i-1}$ und $d_i \in \mathfrak{d}_i$; dies beweist die Summendarstellung (4) für $\mathfrak{i}_k$.

Eine andere Beschreibung der verbindungsirreduziblen Elemente des Verbandes $V(\mathfrak{m})$, die nicht von einer festen Kette (3) ausgeht, wurde in BEHRENS [3], Satz 4, abgeleitet. Wenn nämlich $\mathfrak{m}$ ein distributiver $\mathfrak{R}$-Modul ist, sind für die primären Ringe $\mathfrak{R}_s$ die unitalen $\mathfrak{R}_s$-Moduln $\mathfrak{R}_s\mathfrak{m}=\varepsilon_s\mathfrak{m}$ distributive $\mathfrak{R}_s$-Moduln, deren $\mathfrak{R}_s$-Untermodulverbände aus den Ketten der $\mathfrak{W}_s^a\mathfrak{m}$ bestehen. Die verbindungsirreduziblen $\mathfrak{R}$-Untermoduln von $\mathfrak{m}$ sind dann die Moduln $\mathfrak{R}\mathfrak{W}_s^a\mathfrak{m}$, wobei $1\leqq s\leqq q$ und $0\leqq a\leqq e_s-1$. Dies beweist die erste Behauptung des folgenden Satzes 2, dessen zweite Behauptung die Inklusionen zwischen den verbindungsirreduziblen $\mathfrak{R}$-Untermoduln von $\mathfrak{m}$ und damit die verbandstheoretische Struktur des Verbandes $V(\mathfrak{m})$ innerhalb von $\mathfrak{R}$ beschreibt.

SATZ 2. *Voraussetzungen. 1. Der Artin-Ring $\mathfrak{R}$ mit Einselement sei als additive Gruppe in die primären Ringe $\mathfrak{R}_s$ und eine Untergruppe $\mathfrak{N}$ des Radikales von $\mathfrak{W}$ wie in* (1) *und* (2) *direkt zerlegt. $\mathfrak{W}_s$ sei das Radikal von $\mathfrak{R}_s$ und e_s sein Exponent, $s=1,\ldots,q$. 2. $\mathfrak{m}$ sei ein distributiver $\mathfrak{R}$-K-Modul.*

Behauptungen. 1. Die verbindungsirreduziblen $\mathfrak{R}$-Untermoduln von $\mathfrak{m}$ sind die Moduln $\mathfrak{R}\mathfrak{W}_s^a\mathfrak{m}$, wobei $s=1,\ldots,q$ und $a=0,1,\ldots,e_s-1$ ist. 2. Die Aussagen

$$\mathfrak{R}\mathfrak{W}_s^a\mathfrak{m}\leqq\mathfrak{R}\mathfrak{W}_t^b\mathfrak{m} \tag{7}$$

und

$$\mathfrak{W}_s^{e_s-a-1}\mathfrak{R}\mathfrak{W}_t^b\neq 0 \tag{8}$$

sind miteinander äquivalent.

Ein Korollar hierzu ist der

SATZ 3. *Der Artin-Ring $\mathfrak{R}$ mit Einselement sei distributiv und K-linear darstellbar. Dann ist der $\mathfrak{R}$-Untermodulverband eines distributiven $\mathfrak{R}$-K-Moduls durch $\mathfrak{R}$ bis auf Verbandsisomorphie eindeutig bestimmt.*

BEWEIS DES SATZES 2. Mit den Bezeichnungen von Hilfssatz 1 gilt für die Moduln $\mathfrak{m}_j$ einer maximalen Kette (3) von $\mathfrak{R}$-Untermoduln von $\mathfrak{m}$

$$\varepsilon_s(\mathfrak{m}_j-\mathfrak{m}_{j-1})\neq 0,\quad \text{wobei}\quad s=f(j). \tag{9}$$

Da $\mathfrak{R}_s$ primär, ist der $\mathfrak{R}_s$-Untermodul $\varepsilon_s\mathfrak{m}_j$ von $\varepsilon_s\mathfrak{m}$ gleich $\mathfrak{W}_s^a\mathfrak{m}$ für genau eine Zahl a zwischen 0 und e_s-1. Mit dieser Zahl a gilt auch $\varepsilon_s\mathfrak{m}_{j-1}=\mathfrak{W}_s^{a+1}\mathfrak{m}$, weil der $\mathfrak{R}_s$-Modul (9) irreduzibel und $\neq 0$ ist. Wie in BEHRENS [3], Satz 1, gezeigt, tritt bei diesem Verfahren in einer maximalen Kette (3) von $\mathfrak{R}$-Untermoduln von $\mathfrak{m}$ jede Kombination (s, a) mit $1\leqq s\leqq q$, $0\leqq a\leqq e_s-1$ genau einmal auf. Zu wie in (7) vorgegebenen Paaren (s, a) und (t, b) gibt es daher je eine Zahl j bzw. k zwischen 1 und l mit

$$\varepsilon_s\mathfrak{m}_j=\mathfrak{W}_s^a\mathfrak{m},\quad \varepsilon_t\mathfrak{m}_k=\mathfrak{W}_t^b\mathfrak{m},\quad \varepsilon_s\mathfrak{m}_{j-1}=\mathfrak{W}_s^{a+1}\mathfrak{m},\quad \varepsilon_t\mathfrak{m}_{k+1}=\mathfrak{W}_t^{b+1}\mathfrak{m}. \tag{10}$$

Um hierauf den Hilfssatz 1 anwenden zu können, muß man die Gleichungen

$$\mathfrak{R}\mathfrak{W}_s^a\mathfrak{m}=\mathfrak{R}\mathfrak{d}_j=\mathfrak{i}_j\quad\text{und}\quad\mathfrak{R}\mathfrak{W}_t^b\mathfrak{m}=\mathfrak{R}\mathfrak{d}_k=\mathfrak{i}_k \tag{11}$$

beweisen. Aus (10), aus $\mathfrak{m}_j=\mathfrak{m}_{j-1}+\mathfrak{d}_j$ und aus $\varepsilon_s\mathfrak{d}_j=\mathfrak{d}_j$ folgt $\mathfrak{W}_s^a\mathfrak{m}=\varepsilon_s\mathfrak{m}_j=\varepsilon_s\mathfrak{m}_{j-1}+\mathfrak{d}_j=\mathfrak{W}_s^{a+1}\mathfrak{m}+\mathfrak{d}_j$, also $\mathfrak{W}_s^a\mathfrak{m}=\mathfrak{R}_s\mathfrak{d}_j$, denn die $\mathfrak{R}_s$-Untermoduln von

$V(\mathfrak{m})$ innerhalb des Ringes $\mathfrak{R}$ anzugeben. Dabei legen es die Überlegungen in BEHRENS [2] nahe, folgendermaßen vorzugehen:

HILFSSATZ 1. *Sei*

$$0 = \mathfrak{m}_0 < \mathfrak{m}_1 < \cdots < \mathfrak{m}_l = \mathfrak{m} \tag{3}$$

eine maximale Kette von $\mathfrak{R}$-Untermoduln des distributiven $\mathfrak{R}-K$-Moduls $\mathfrak{m}$. Jeder der $\mathfrak{R}$-irreduziblen Faktormoduln $\mathfrak{m}_i - \mathfrak{m}_{i-1}$ stellt demnach genau einen der einfachen Ringe $\overline{\mathfrak{R}}_s$ aus (2) treu dar, etwa $\overline{\mathfrak{R}}_r$, und annulliert die anderen. Setze $r = f(i)$. Zu jedem $i = 1, \ldots, l$ existiert eine additive Untergruppe $\mathfrak{d}_i$ von $\mathfrak{m}$ mit $\varepsilon_r \mathfrak{d}_i = \mathfrak{d}_i$ derart, daß $\mathfrak{m}_i$ als Gruppe die direkte Summe von $\mathfrak{m}_{i-1}$ und $\mathfrak{d}_i$ ist. Setze $c_{ik} = 0$, wenn $\mathfrak{d}_i \cap \mathfrak{R}\mathfrak{d}_k = 0$ ist, und andernfalls $c_{ik} = 1$:

Behauptung.

$$\mathfrak{i}_k = \sum_{c_{ik}=1} \mathfrak{d}_i = \mathfrak{R}\mathfrak{d}_k, \qquad k = 1, \ldots, l, \tag{4}$$

sind die verbindungsirreduziblen $\mathfrak{R}$-Untermoduln von $\mathfrak{m}$. Bezüglich ihrer Inklusionen gilt:

$$\mathfrak{i}_j \subseteqq \mathfrak{i}_k \quad \textit{ist äquivalent} \quad c_{jk} = 1. \tag{5}$$

BEWEIS. Die Existenz der obigen $\mathfrak{d}_i$ in $\mathfrak{m} = \mathfrak{d}_1 + \cdots + \mathfrak{d}_l$ folgt daraus, daß $\mathfrak{m}_i - \mathfrak{m}_{i-1}$ ein $K^{(r)}$-Rechtsmodul ist: Seien $b_1, \ldots, b_d$ eine $K^{(r)}$-Basis, wobei alle $b_m \in \mathfrak{m}_i$. Dann enthält $\mathfrak{m}_i$ die als Gruppe zu $\mathfrak{m}_i - \mathfrak{m}_{i-1}$ isomorphe Gruppe $b_1 K^{(r)} + \cdots + b_d K^{(r)} = \mathfrak{d}_i'$. Also ist $\mathfrak{m}_i = \mathfrak{m}_{i-1} + \mathfrak{d}_i'$. Da ε_r den Modul $\mathfrak{m}_i - \mathfrak{m}_{i-1}$ elementweise reproduziert, ist $(1 - \varepsilon_r)\mathfrak{d}_i' \subseteqq \mathfrak{m}_{i-1}$ und demnach $\mathfrak{d}_i = \varepsilon_r \mathfrak{d}_i'$ als Gruppe direkter Summand in $\mathfrak{m}_i = \mathfrak{m}_{i-1} + \mathfrak{d}_i$ mit $\varepsilon_r \mathfrak{d}_i = \mathfrak{d}_i$. — Nach BEHRENS [3], S. 268–269, läßt sich jeder $\mathfrak{R}$-Untermodul $\mathfrak{u}$ von $\mathfrak{m}$ durch die als additive Gruppe direkte Summe

$$\mathfrak{u} = \sum_{i=1}^{q} \mathfrak{W}_i^{u_i} \mathfrak{m} \tag{6}$$

mit eindeutig bestimmten Exponenten u_i zwischen 0 und e_i ausdrücken. Analog seien die Zahlen v_i die Exponenten von $\mathfrak{v} \in V(\mathfrak{m})$. Dann gehören zu der Verbindung $\mathfrak{u} \cup \mathfrak{v}$ die Exponenten $s_i = \operatorname{Min}(u_i, v_i)$. Angenommen nun, $\mathfrak{i}_k = \mathfrak{R}\mathfrak{d}_k$ wäre reduzibel, also $\mathfrak{i}_k = \mathfrak{u} \cup \mathfrak{v}$, aber $\mathfrak{i}_k$ von $\mathfrak{u}$ und $\mathfrak{v}$ verschieden. Mit der Bezeichnung $s = f(k)$ ist $\mathfrak{i}_k = \mathfrak{R}\mathfrak{d}_k \subseteqq \mathfrak{u}$ äquivalent mit $\mathfrak{d}_k = \varepsilon_s \mathfrak{d}_k \subseteqq \mathfrak{u}$, also auch mit $\mathfrak{d}_k \subseteqq \varepsilon_s \mathfrak{u} = \mathfrak{W}_s^{u_s} \mathfrak{m}$, denn da $\mathfrak{R}_s$ ein primärer Ring ist, bilden die $\mathfrak{R}_s$-Untermoduln von $\mathfrak{m}$ die Kette der $\mathfrak{W}_s^b \mathfrak{m}$, $0 \leqq b \leqq e_s$. Analog für $\mathfrak{v}$ und für $\mathfrak{u} \cup \mathfrak{v}$. Aus $\mathfrak{i}_k = \mathfrak{u} \cup \mathfrak{v}$ folgt demnach $\mathfrak{d}_k \subseteqq \mathfrak{W}_s^m \mathfrak{m}$, wobei $m = \min(u_s, v_s)$ ist. Aus $\mathfrak{i}_k \subsetneqq \mathfrak{u}$ und $\mathfrak{i}_k \subsetneqq \mathfrak{v}$ folgt andererseits $\mathfrak{d}_k \subsetneqq \mathfrak{W}_s^{u_s} \mathfrak{m}$ und $\mathfrak{d}_k \subsetneqq \mathfrak{W}_s^{v_s} \mathfrak{m}$, q.e.a. — Dies impliziert $\mathfrak{m}_i = \mathfrak{i}_1 \cup \cdots \cup \mathfrak{i}_i$. Sei x_i ein nicht in $\mathfrak{m}_{i-1}$ enthaltenes Element aus $\mathfrak{m}_i \cap \mathfrak{R}\mathfrak{d}_k$. Da auch der $\mathfrak{R}$-Modul $\mathfrak{R}x_i$ in diesem Durchschnitt liegt, tritt in dessen Darstellung als Verbindung von $\cup$-irreduziblen Elementen $\mathfrak{i}_i$ auf. Aus $\mathfrak{R}x_i \subseteqq \mathfrak{R}\mathfrak{d}_k$ folgt dann $\mathfrak{d}_i \subseteqq \mathfrak{R}\mathfrak{d}_k$. (Insbesondere ist $c_{ik} = 1$ äquivalent $\mathfrak{d}_i \subseteqq \mathfrak{R}\mathfrak{d}_k$.) Demnach enthält schließlich $\mathfrak{R}\mathfrak{d}_k$ mit $m_{i-1} + d_i$ auch die Elemente $m_{i-1} \in \mathfrak{m}_{i-1}$ und $d_i \in \mathfrak{d}_i$; dies beweist die Summendarstellung (4) für $\mathfrak{i}_k$.

Eine andere Beschreibung der verbindungsirreduziblen Elemente des Verbandes $V(\mathfrak{m})$, die nicht von einer festen Kette (3) ausgeht, wurde in BEHRENS [3], Satz 4, abgeleitet. Wenn nämlich $\mathfrak{m}$ ein distributiver $\mathfrak{R}$-Modul ist, sind für die primären Ringe $\mathfrak{R}_s$ die unitalen $\mathfrak{R}_s$-Moduln $\mathfrak{R}_s\mathfrak{m} = \varepsilon_s\mathfrak{m}$ distributive $\mathfrak{R}_s$-Moduln, deren $\mathfrak{R}_s$-Untermodulverbände aus den Ketten der $\mathfrak{W}_s^a\mathfrak{m}$ bestehen. Die verbindungsirreduziblen $\mathfrak{R}$-Untermoduln von $\mathfrak{m}$ sind dann die Moduln $\mathfrak{R}\mathfrak{W}_s^a\mathfrak{m}$, wobei $1 \leqq s \leqq q$ und $0 \leqq a \leqq e_s - 1$. Dies beweist die erste Behauptung des folgenden Satzes 2, dessen zweite Behauptung die Inklusionen zwischen den verbindungsirreduziblen $\mathfrak{R}$-Untermoduln von $\mathfrak{m}$ und damit die verbandstheoretische Struktur des Verbandes $V(\mathfrak{m})$ innerhalb von $\mathfrak{R}$ beschreibt.

SATZ 2. *Voraussetzungen. 1. Der Artin-Ring $\mathfrak{R}$ mit Einselement sei als additive Gruppe in die primären Ringe $\mathfrak{R}_s$ und eine Untergruppe $\mathfrak{N}$ des Radikales von $\mathfrak{W}$ wie in* (1) *und* (2) *direkt zerlegt. $\mathfrak{W}_s$ sei das Radikal von $\mathfrak{R}_s$ und e_s sein Exponent, $s = 1, \ldots, q$. 2. $\mathfrak{m}$ sei ein distributiver $\mathfrak{R}$-K-Modul.*

Behauptungen. 1. Die verbindungsirreduziblen $\mathfrak{R}$-Untermoduln von $\mathfrak{m}$ sind die Moduln $\mathfrak{R}\mathfrak{W}_s^a\mathfrak{m}$, wobei $s = 1, \ldots, q$ und $a = 0, 1, \ldots, e_s - 1$ ist. 2. Die Aussagen

$$\mathfrak{R}\mathfrak{W}_s^a\mathfrak{m} \leqq \mathfrak{R}\mathfrak{W}_t^b\mathfrak{m} \tag{7}$$

und

$$\mathfrak{W}_s^{e_s-a-1}\mathfrak{R}\mathfrak{W}_t^b \neq 0 \tag{8}$$

sind miteinander äquivalent.

Ein Korollar hierzu ist der

SATZ 3. *Der Artin-Ring $\mathfrak{R}$ mit Einselement sei distributiv und K-linear darstellbar. Dann ist der $\mathfrak{R}$-Untermodulverband eines distributiven $\mathfrak{R}$-K-Moduls durch $\mathfrak{R}$ bis auf Verbandsisomorphie eindeutig bestimmt.*

BEWEIS DES SATZES 2. Mit den Bezeichnungen von Hilfssatz 1 gilt für die Moduln $\mathfrak{m}_j$ einer maximalen Kette (3) von $\mathfrak{R}$-Untermoduln von $\mathfrak{m}$

$$\varepsilon_s(\mathfrak{m}_j - \mathfrak{m}_{j-1}) \neq 0, \quad \text{wobei} \quad s = f(j). \tag{9}$$

Da $\mathfrak{R}_s$ primär, ist der $\mathfrak{R}_s$-Untermodul $\varepsilon_s\mathfrak{m}_j$ von $\varepsilon_s\mathfrak{m}$ gleich $\mathfrak{W}_s^a\mathfrak{m}$ für genau eine Zahl a zwischen 0 und $e_s - 1$. Mit dieser Zahl a gilt auch $\varepsilon_s\mathfrak{m}_{j-1} = \mathfrak{W}_s^{a+1}\mathfrak{m}$, weil der $\mathfrak{R}_s$-Modul (9) irreduzibel und $\neq 0$ ist. Wie in BEHRENS [3], Satz 1, gezeigt, tritt bei diesem Verfahren in einer maximalen Kette (3) von $\mathfrak{R}$-Untermoduln von $\mathfrak{m}$ jede Kombination (s, a) mit $1 \leqq s \leqq q$, $0 \leqq a \leqq e_s - 1$ genau einmal auf. Zu wie in (7) vorgegebenen Paaren (s, a) und (t, b) gibt es daher je eine Zahl j bzw. k zwischen 1 und l mit

$$\varepsilon_s\mathfrak{m}_j = \mathfrak{W}_s^a\mathfrak{m}, \quad \varepsilon_t\mathfrak{m}_k = \mathfrak{W}_t^b\mathfrak{m}, \quad \varepsilon_s\mathfrak{m}_{j-1} = \mathfrak{W}_s^{a+1}\mathfrak{m}, \quad \varepsilon_t\mathfrak{m}_{k+1} = \mathfrak{W}_t^{b+1}\mathfrak{m}. \tag{10}$$

Um hierauf den Hilfssatz 1 anwenden zu können, muß man die Gleichungen

$$\mathfrak{R}\mathfrak{W}_s^a\mathfrak{m} = \mathfrak{R}\mathfrak{d}_j = \mathfrak{i}_j \quad \text{und} \quad \mathfrak{R}\mathfrak{W}_t^b\mathfrak{m} = \mathfrak{R}\mathfrak{d}_k = \mathfrak{i}_k \tag{11}$$

beweisen. Aus (10), aus $\mathfrak{m}_j = \mathfrak{m}_{j-1} + \mathfrak{d}_j$ und aus $\varepsilon_s\mathfrak{d}_j = \mathfrak{d}_j$ folgt $\mathfrak{W}_s^a\mathfrak{m} = \varepsilon_s\mathfrak{m}_j = \varepsilon_s\mathfrak{m}_{j-1} + \mathfrak{d}_j = \mathfrak{W}_s^{a+1}\mathfrak{m} + \mathfrak{d}_j$, also $\mathfrak{W}_s^a\mathfrak{m} = \mathfrak{R}_s\mathfrak{d}_j$, denn die $\mathfrak{R}_s$-Untermoduln von

$\varepsilon_s\mathfrak{m}$ bilden eine Kette. Dies zieht wegen $\mathfrak{R}\,\mathfrak{R}_s\mathfrak{b}_j=\mathfrak{R}\mathfrak{b}_j$ nach Hilfssatz 1 die ersten Gleichungen in (11) nach sich. Analog beweist man die zweiten. — Natürlich gilt

$$\mathfrak{W}_s^{e_s-a-1}\mathfrak{b}_j\neq 0, \tag{12}$$

denn sonst wäre wegen (11), bereits $\mathfrak{W}_s^{e_s-1}\mathfrak{m}=0$ entgegen der Treue von $\mathfrak{m}$. Die in Satz 2 genannte Aussage (7) impliziert wegen (11) die Relation

$$\mathfrak{b}_j\subseteqq\mathfrak{R}\,\mathfrak{b}_k=\mathfrak{R}\,\mathfrak{W}_i^b\,\mathfrak{m}, \tag{13}$$

also wegen (12) die Behauptung (8) des Satzes 2.

Sei nun umgekehrt (8) erfüllt. Für genau ein k ist dann $\mathfrak{u}=\mathfrak{W}_s^{e_s-a-1}\mathfrak{R}\,\mathfrak{W}_i^b\mathfrak{m}=\mathfrak{W}_s^{e_s-a-1}\mathfrak{i}_k=\mathfrak{W}_s^{e_s-a-1}\mathfrak{R}\,\mathfrak{b}_k$ ein von 0 verschiedener, in $\varepsilon_s\mathfrak{R}\,\mathfrak{b}_k$ enthaltener $\mathfrak{R}_s$-Untermodul von $\varepsilon_s\mathfrak{m}$. Hierbei ist $\varepsilon_s\mathfrak{R}\,\mathfrak{b}_k=\mathfrak{W}_s^p\mathfrak{m}$ für einen geeigneten Exponenten p zwischen 0 und e_s-1. Aus $\mathfrak{u}\neq 0$ folgt, daß $\mathfrak{W}_s^{e_s-a-1}\mathfrak{W}_s^p\mathfrak{m}\neq 0$, also wegen, $\mathfrak{W}_s^{e_s}=0$ die Ungleichung $e_s-a-1+p\leqq e_s-1$ und damit $p\leqq a$ gilt. Die damit bewiesene Inklusion $\mathfrak{W}_s^a\mathfrak{m}\subseteqq\mathfrak{W}_s^p\mathfrak{m}=\varepsilon_s\mathfrak{R}\,\mathfrak{b}_k$ zieht wegen $\varepsilon_s\mathfrak{b}_k=\mathfrak{b}_k$ nach sich $\mathfrak{R}\,\mathfrak{W}_s^a\mathfrak{m}\subseteqq\mathfrak{R}\,\mathfrak{b}_k\subseteqq\mathfrak{R}\,\mathfrak{W}_i^b\mathfrak{m}$, also (7). Q.e.d.

Die Relationen (8) haben etwas mit der direkten Zerlegung von $\mathfrak{R}$ in zweiseitige Ideale mit Hilfe von Blöcken zu tun. (ARTIN, NESBITT, THRALL [1], IX.6).

§ 2

Ein primärer, distributiv und K-linear darstellbarer Ring besitzt nur ein einziges minimales Ideal, nämlich $\mathfrak{W}^{e-1}$, wie in Teil I ,Satz 3, bewiesen wurde. Auch ein beliebiger distributiv und K-linear darstellbarer Artin-Ring enthält nur wenige Ideale, nämlich die im folgenden Satz 5 genannten.

Wieder sei $\mathfrak{R}$ durch primäre Ringe $\mathfrak{R}_i$ und eine additive Gruppe $\mathfrak{N}_i$ aus dem Radikal $\mathfrak{W}$ in der Form

$$\mathfrak{R}=\mathfrak{R}_1+\cdots+\mathfrak{R}_q+\mathfrak{N} \tag{1}$$

dargestellt, wobei

$$\mathfrak{R}_i=\varepsilon_i\mathfrak{R}\,\varepsilon_i,\qquad \mathfrak{N}=\sum_{i\neq k}\varepsilon_i\,\mathfrak{R}\,\varepsilon_k,\qquad \varepsilon_1+\cdots+\varepsilon_q=1,\ \varepsilon_i$$

orthogonale Idempotente sind. Mit $\mathfrak{W}_i$ sei das Radikal des Ringes $\mathfrak{R}_i$ bezeichnet, und e_i sei sein Exponent.

Die Aufstellung der minimalen Ideale in $\mathfrak{R}$ wird vorbereitet durch den

SATZ 4. *$\mathfrak{m}$ sei ein distributiver $\mathfrak{R}$-K-Modul, $\mathfrak{M}$ ein minimales Ideal in $\mathfrak{R}$, dann ist $\mathfrak{M}\mathfrak{m}$ ein minimaler $\mathfrak{R}$-Untermodul von $\mathfrak{m}$, und es existiert genau ein ε_i unter den $\varepsilon_1,\ldots,\varepsilon_q$, für welches $\varepsilon_i\mathfrak{M}=\mathfrak{M}$ ist.*

BEWEIS. $\mathfrak{W}\mathfrak{M}=0$ weil sonst $\mathfrak{M}=\mathfrak{W}\mathfrak{M}=\mathfrak{W}^2\mathfrak{M}=\cdots=0$ wäre. Also ist $\mathfrak{M}\mathfrak{m}$ ein $\mathfrak{R}/\mathfrak{W}$-Modul, also voll reduzibel. Angenommen $\mathfrak{M}\mathfrak{m}$ wäre reduzibel. Dann ist $\mathfrak{M}\mathfrak{m}=\mathfrak{u}+\mathfrak{v}+\cdots$, wobei $\mathfrak{u}$ und $\mathfrak{v}$ irreduzible $\mathfrak{R}$-Moduln sind und die Summe als $\mathfrak{R}$-Modul direkt ist. $\mathfrak{u}$ und $\mathfrak{v}$ können nicht zueinander $\mathfrak{R}$-isomorph sein, weil dies, wie im Beweis von Satz 1 erläutert wurde, der Distributivität von $V(\mathfrak{m})$ widerspräche. Sei etwa $\varepsilon_1\mathfrak{u}=\mathfrak{u}$, $\varepsilon_2\mathfrak{v}=\mathfrak{v}$. Auch $\varepsilon_1\mathfrak{M}$ ist Ideal in $\mathfrak{R}$,

denn jedes Element $\varrho \in \mathfrak{R}$ läßt sich nach (1) in der Form $\varrho = \sum \varepsilon_i \varrho \varepsilon_i + \nu$, $\nu \in \mathfrak{W}$ schreiben, also ist $\varrho \varepsilon_1 \mathfrak{M} = \varepsilon_1 \varrho \varepsilon_1 \mathfrak{M}$, da $\mathfrak{W}\mathfrak{M} = 0$. Das Ideal $\varepsilon_1 \mathfrak{M}$ ist von 0 verschieden, weil $\mathfrak{u} = \varepsilon_1 \mathfrak{u} \subseteq \mathfrak{M}\mathfrak{m}$. Daher ist $\varepsilon_1 \mathfrak{M} = \mathfrak{M}$. Dies widerspricht aber $\varepsilon_2 \mathfrak{v} = \mathfrak{v} \subseteq \mathfrak{M}\mathfrak{m} = \varepsilon_1 \mathfrak{M}$. Schließlich ist für $j \neq 1$ das Produkt $\varepsilon_j \mathfrak{M} = \varepsilon_j \varepsilon_1 \mathfrak{M} = 0$.

Wenn $\mathfrak{m}$ ein distributiver $\mathfrak{R}$-K-Modul ist, dann sind nach Satz 1 alle $\mathfrak{R}$-Untermoduln und damit auch alle Faktormoduln dieser Untermoduln von $\mathfrak{m}$ ebenfalls $\mathfrak{R}$-K-Moduln. Dies gilt insbesondere für die $\mathfrak{R}$-irreduziblen unter diesen Faktormoduln. Sei etwa $\bar{\mathfrak{u}}$ ein solcher Modul, also ein irreduzibler $\overline{\mathfrak{R}}_i$-Modul für passendes i. Die Elemente aus K induzieren dann $\overline{\mathfrak{R}}_i$-Endomorphismen von $\bar{\mathfrak{u}}$, also Elemente aus dem Zentralisator $K^{(i)}$ des $\overline{\mathfrak{R}}_i$-Mod $\bar{\mathfrak{u}}$

Satz 5. *Voraussetzung. Der Ring $\mathfrak{R}$ sei distributiv und K-linear darstellbar. Hierbei sei der Ring K so groß, daß er für alle $\mathfrak{R}$-irreduziblen Faktormoduln der $\mathfrak{R}$-Untermoduln eines distributiven $\mathfrak{R}$-K-Moduls $\mathfrak{m}$ deren volle Zentralisatoren $K^{(i)}$ induziert.*

Behauptung. Die Untermenge

$$\mathfrak{M}_{ik} = \{\tau \mid \tau \in \mathfrak{R},\ \varepsilon_i \tau \varepsilon_k = \tau,\ \mathfrak{W}_j \tau = 0 = \tau \mathfrak{W}_k\} \tag{2}$$

von $\mathfrak{R}$ ist dann und nur dann minimales Ideal in $\mathfrak{R}$, wenn $\mathfrak{M}_{ik} \neq 0$ ist und in $\mathfrak{R}$ die Beziehungen

$$(1 - \varepsilon_i)\, \mathfrak{R}\, \mathfrak{W}_i^{e_i - 1} = 0 = \mathfrak{W}_k^{e_k - 1}\, \mathfrak{R}\, (1 - \varepsilon_k) \tag{3}$$

gelten. — Jedes minimale Ideal aus $\mathfrak{R}$ hat die Gestalt (2) mit (3).

Zusatz. *Die Gln. (3) bedeuten für jeden distributiven $\mathfrak{R}$-K-Modul $\mathfrak{m}$, daß unter allen verbindungsirreduziblen $\mathfrak{R}$-Untermoduln von $\mathfrak{m}$ die Moduln $\mathfrak{R}\mathfrak{W}_i^{e_i - 1}\mathfrak{m}$ und $\mathfrak{R}\varepsilon_k\mathfrak{m}$ minimal bzw. maximal sind*

Beweis des Zusatzes. Nach Satz 2 ist $\mathfrak{R}\mathfrak{W}_i^{e_i-1}\mathfrak{m}$ genau dann minimal, wenn für alle $(s, a) \neq (i, e_i - 1)$ gilt $\mathfrak{W}_s^{e_s - a - 1}\mathfrak{R}\mathfrak{W}_i^{e_i - 1} = 0$ oder, was auf dasselbe hinausläuft, wenn für alle $s \neq i$ gilt $0 = \mathfrak{W}_s^0 \mathfrak{R}\mathfrak{W}_i^{e_i-1} = \varepsilon_s \mathfrak{R}\mathfrak{W}_i^{e_i-1}$. Dies ist wegen $\varepsilon_1 + \cdots + \varepsilon_q = 1$ mit der ersten Gleichung in (3) äquivalent. Ähnlich ist $\mathfrak{R}\varepsilon_k\mathfrak{m} = \mathfrak{R}\mathfrak{W}_k^0\mathfrak{m}$ gemäß Satz 2 genau dann maximal, wenn die zweite Gleichung in (3) erfüllt ist.

Beweis des Satzes 5. 1. $\mathfrak{M}$ sei ein minimales Ideal in $\mathfrak{R}$ und $\mathfrak{m}$ ein distributiver $\mathfrak{R}$-K-Modul. Nach Satz 4 ist $\mathfrak{M}\mathfrak{m}$ ein minimaler $\mathfrak{R}$-Modul. Für ihn existiert genau ein ε_i unter den $\varepsilon_1, \ldots, \varepsilon_q$, für welches $\varepsilon_i \mathfrak{M} = \mathfrak{M}$ ist. Dies zieht $\varepsilon_i \tau = \tau$ für alle $\tau \in \mathfrak{M}$ nach sich. Als $\mathfrak{R}_i$-Modul ist $\varepsilon_i \mathfrak{M}\mathfrak{m} = \mathfrak{W}_i^a \mathfrak{m}$, also $\mathfrak{M}\mathfrak{m} = \mathfrak{R}\mathfrak{M}\mathfrak{m} = \mathfrak{R}\varepsilon_i\mathfrak{M}\mathfrak{m} = \mathfrak{R}\mathfrak{W}_i^a\mathfrak{m}$. Da $\mathfrak{M}\mathfrak{m}$ minimaler $\mathfrak{R}$-Modul, also erst recht minimal unter den verbindungsirreduziblen $\mathfrak{R}$-Untermoduln von $\mathfrak{m}$ ist, ist $a = e_i - 1$ und, nach dem Zusatz, die erste Gleichung in (3) richtig. Ferner impliziert die Minimalität des Ideales $\mathfrak{M}$ die Gleichungen $0 = \mathfrak{W}\mathfrak{M} = \mathfrak{M}\mathfrak{W}$, also auch $\mathfrak{W}_i \tau = 0 = \tau \mathfrak{W}_k$ für alle $\tau \in \mathfrak{M}$. Durch den minimalen $\mathfrak{R}$-Untermodul $\mathfrak{m}_1 = \mathfrak{M}\mathfrak{m}$ von $\mathfrak{m}$ sei nun eine maximale Kette

$$0 = \mathfrak{m}_0 \subset \mathfrak{m}_1 \subset \cdots \subset \mathfrak{m}_l = \mathfrak{m} \tag{4}$$

von $\mathfrak{R}$-Untermoduln von $\mathfrak{m}$ gelegt. Wegen $\varepsilon_1+\cdots+\varepsilon_q=1$ ist für mindestens ein ε_k das Produkt $\mathfrak{M}\varepsilon_k\neq 0$. Angenommen es wäre auch $\mathfrak{M}\varepsilon_p\neq 0$ für ein $p\neq k$. Dann existiert ein Zahlenpaar (u, v) mit

$$(5)\qquad \begin{cases} \mathfrak{M}\varepsilon_k\mathfrak{m}_u\neq 0, & \mathfrak{M}\varepsilon_p\mathfrak{m}_v\neq 0 \\ \mathfrak{M}\varepsilon_k\mathfrak{m}_{u-1}=0, & \mathfrak{M}\varepsilon_p\mathfrak{m}_{v-1}=0. \end{cases}$$

Da $\mathfrak{m}_u-\mathfrak{m}_{u-1}$ und $\mathfrak{m}_v-\mathfrak{m}_{v-1}$ irreduzible $\mathfrak{R}$-Moduln sind, aber $\varepsilon_k\neq\varepsilon_p$ ist, ist $u\neq v$. Sei etwa $u<v$. Dann ist das minimale Ideal $\mathfrak{M}=\mathfrak{M}\varepsilon_k\mathfrak{R}=\mathfrak{M}\varepsilon_p\mathfrak{R}$, also $0\neq\mathfrak{M}\varepsilon_k\mathfrak{m}_u=\mathfrak{M}\varepsilon_k\mathfrak{R}\mathfrak{m}_u=\mathfrak{M}\varepsilon_p\mathfrak{R}\mathfrak{m}_u\subseteq\mathfrak{M}\varepsilon_p\mathfrak{R}\mathfrak{m}_{v-1}=\mathfrak{M}\varepsilon_p\mathfrak{m}_{v-1}=0$. Q.e.a. Dies beweist

$$(6)\qquad \mathfrak{M}\varepsilon_p=0 \quad \text{für} \quad p\neq k, \qquad \mathfrak{M}=\mathfrak{M}\varepsilon_k$$

und damit $\tau\varepsilon_k=\tau$ für alle $\tau\in\mathfrak{M}$. Nach (5) ist nun $\varepsilon_k\mathfrak{m}_u\neq 0$, also als $\mathfrak{R}_k$-Modul $0\neq\varepsilon_k\mathfrak{m}_u=\mathfrak{W}_k^a\mathfrak{m}$. Hierbei ist $a=0$, denn aus $0\neq\mathfrak{M}\varepsilon_k\mathfrak{m}_u=\mathfrak{M}\mathfrak{W}_k^a\mathfrak{m}$ folgt $\mathfrak{M}\mathfrak{W}_k^a\mathfrak{R}=\mathfrak{M}=\mathfrak{M}\varepsilon_k$, also $\mathfrak{M}=\mathfrak{M}\mathfrak{W}_k^a$ und damit $a=0$, denn $\mathfrak{W}_k$ ist nilpotent. Die damit bewiesene Gleichung

$$(7)\qquad \varepsilon_k\mathfrak{m}_u=\varepsilon_k\mathfrak{m}$$

kann man zum Beweis der zweiten Gleichung in (3) heranziehen: Wie in Hilfssatz 1 seien die $\mathfrak{d}_j$-Untergruppen von $\mathfrak{m}$ mit $\mathfrak{m}_j=\mathfrak{m}_{j-1}+\mathfrak{d}_j$, wobei $\mathfrak{d}_j$ von genau einem der Idempotente $\varepsilon_1, \ldots, \varepsilon_q$ reproduziert und von den anderen annulliert wird. Wegen (7) ist $\varepsilon_k(\mathfrak{d}_{u+1}+\cdots+\mathfrak{d}_l)=0$. Nach Hilfssatz 1 ist $\mathfrak{i}_u=\mathfrak{R}\mathfrak{d}_u$ genau dann maximal unter den verbindungsirreduziblen $\mathfrak{R}$-Untermoduln von $\mathfrak{m}$, wenn $c_{u,u+1}=\cdots=c_{ul}=0$ ist. Angenommen es wäre $c_{ur}=1$ für ein $r>u$, also $\mathfrak{d}_u\cap\mathfrak{R}\mathfrak{d}_r\neq 0$. Genauer würde dann $\mathfrak{d}_u\subseteq\mathfrak{R}\mathfrak{d}_r$ gelten, wie in §1 beim Beweis des Hilfsatzes 1 gezeigt wurde. Wegen (5) wäre $0\neq\mathfrak{M}\varepsilon_k\mathfrak{d}_u=\mathfrak{M}\mathfrak{d}_u\subseteq\mathfrak{M}\mathfrak{R}\mathfrak{d}_r=\mathfrak{M}\varepsilon_k\mathfrak{d}_r$, also $\varepsilon_k\mathfrak{d}_r\neq 0$, q.e.a. Dies impliziert die Maximalität von $\mathfrak{i}_u=\mathfrak{R}\mathfrak{d}_u$ unter den verbindungsirreduziblen $\mathfrak{R}$-Untermoduln von $\mathfrak{m}$. Wegen (5) und (7) ist $\mathfrak{d}_u=\varepsilon_k\mathfrak{d}_u\subseteq\varepsilon_k\mathfrak{m}$, also $\mathfrak{R}\mathfrak{d}_u=\mathfrak{R}\varepsilon_k\mathfrak{m}$, denn auch $\mathfrak{R}\varepsilon_k\mathfrak{m}$ ist verbindungsirreduzibel. Dies zieht auf Grund des bereits bewiesenen Zusatzes die zweite Gleichung in (3) nach sich.

2. Zum Beweis der Umkehrung sei $\mathfrak{M}=\mathfrak{M}_{ik}\neq 0$ eine Menge (2). Außerdem gelte (3) für den Ring $\mathfrak{R}$. Daß $\mathfrak{M}$ ein Linksideal ist, kann man folgendermaßen zeigen: Jedes Element $\varrho\in\mathfrak{R}$ läßt sich nach (1) in der Form

$$(8)\qquad \varrho=\sum_{1\leq j\leq q}\varepsilon_j\varrho_j+\nu, \qquad \nu\in\mathfrak{N}=\sum_{j\neq k}\varepsilon_j\mathfrak{R}\varepsilon_k,$$

schreiben. Wegen $\mathfrak{M}=\varepsilon_i\mathfrak{M}$ ist

$$(9)\qquad \varrho\mathfrak{M}=\varepsilon_i\varrho\varepsilon_i\mathfrak{M}+\nu\mathfrak{M}.$$

Sei $\mathfrak{m}$ ein distributiver $\mathfrak{R}$-K-Modul, wie in der Voraussetzung. Sein $\mathfrak{R}_i$-Untermodul $\varepsilon_i\mathfrak{M}\mathfrak{m}=\mathfrak{M}\mathfrak{m}$ ist gleich $\mathfrak{W}_i^a\mathfrak{m}$ mit $a=e_i-1$, da $\mathfrak{W}_i\mathfrak{m}=0$. Wäre nun in (9) das Produkt $\nu\mathfrak{M}\neq 0$, dann wäre, wegen der Treue des $\mathfrak{R}$-Moduls $\mathfrak{m}$,

$$0\neq\sum_{j,\,j\neq i}\varepsilon_j\mathfrak{R}\varepsilon_i\mathfrak{M}\mathfrak{m}=(1-\varepsilon_i)\mathfrak{R}\mathfrak{W}_i^{e_i-1}\mathfrak{m},$$

im Widerspruch zur ersten Gleichung in (3). Gemäß (9) ist daher $\mathfrak{M}$ sicher dann Linksideal in $\mathfrak{R}$, wenn es ein $\mathfrak{R}_i$-Linksmodul ist. Letzteres folgt für $\varrho \in \mathfrak{R}_i$, $\tau \in \mathfrak{M}$ aus (2), das heißt aus den Gleichungen $\varepsilon_i \tau \varepsilon_k = \tau$, $\mathfrak{W}_i \tau = 0 = \tau \mathfrak{W}_k$, da nämlich $\varepsilon_i \varrho \tau \varepsilon_k = \varepsilon_i \varrho \varepsilon_i \tau \varepsilon_k = \varrho \tau$, $\mathfrak{W}_i \varrho \tau \leqq \mathfrak{W}_i \tau = 0 = \varrho \tau \mathfrak{W}_k$. — Da $\mathfrak{M}$ Linksideal ist, ist $\mathfrak{M}\mathfrak{m} = \mathfrak{R}\mathfrak{M}\mathfrak{m} = \mathfrak{R}\mathfrak{W}_i^{e-1}\mathfrak{m}$ nach (3) und Satz 2 ein minimaler $\mathfrak{R}$-Modul. Durch $\mathfrak{m}_1 = \mathfrak{M}\mathfrak{m}$ sei eine maximale Kette (4) von $\mathfrak{R}$-Untermoduln von $\mathfrak{m}$ gelegt. In ihr existiert ein Index u mit $\varepsilon_k \mathfrak{m}_u = \varepsilon_k \mathfrak{m}$, aber $\mathfrak{W}_k \varepsilon_k \mathfrak{m} = \mathfrak{W}_k \varepsilon_k \mathfrak{m}_u = \varepsilon_k \mathfrak{m}_{u-1}$. Für dieses u gelten $\mathfrak{M}\mathfrak{m} = \mathfrak{M}\varepsilon_k \mathfrak{m} = \mathfrak{M}\mathfrak{m}_u$ und $0 = \mathfrak{M}\mathfrak{W}_k \mathfrak{m} = \mathfrak{M}\mathfrak{W}_k \mathfrak{m}_u = \mathfrak{M}\mathfrak{m}_{u-1}$. Führt man wieder die $\mathfrak{d}_j$ mit $\mathfrak{m}_j = \mathfrak{m}_{j-1} + \mathfrak{d}_j$ aus Hilfssatz 1 ein, so annulliert wegen $\mathfrak{M} = \mathfrak{M}\varepsilon_k$, $\varepsilon_k \mathfrak{m} = \varepsilon_k \mathfrak{m}_u$, $\mathfrak{M}\mathfrak{m}_{u-1} = 0$, die Menge $\mathfrak{M}$ alle $\mathfrak{m}_j$ außer $\mathfrak{d}_u$. Da $\mathfrak{R}\mathfrak{m}_{u-1} = \mathfrak{m}_{u-1}$, gilt auch $\mathfrak{M}\mathfrak{R}\mathfrak{d}_j = 0$ für alle $j < u$. Dies ist auch für alle $j > u$ richtig. Beweis: Auf Grund des Zusatzes zu Satz 5 bedeutet die zweite Gleichung in (3), daß $\mathfrak{R}\varepsilon_k \mathfrak{m}_u = \mathfrak{R}\varepsilon_k \mathfrak{m}$ maximal unter den verbindungsirreduziblen $\mathfrak{R}$-Untermoduln von $\mathfrak{m}$ ist, also nach Hilfssatz 1 die Gleichungen $c_{u,u+1} = \cdots = c_{ul} = 0$. Daher tritt für $j > u$, ebenfalls nach Hilfssatz 1, $\mathfrak{d}_u$ nicht unter den $\mathfrak{d}_h$ in $\mathfrak{R}\mathfrak{d}_j = \sum_{c_{hj}=1} \mathfrak{d}_h$ auf. Wegen $\mathfrak{M}\mathfrak{d}_j = 0$ für $j \neq u$, ist $\mathfrak{M}\mathfrak{R}\mathfrak{d}_j = \sum \mathfrak{M}\mathfrak{d}_h = 0$. Folglich ist, wie behauptet, $\mathfrak{M}\mathfrak{R}\mathfrak{d}_j = 0$ für alle $j \neq u$. Hieraus läßt sich ableiten, daß $\mathfrak{M}$ ein Rechtsideal ist. Beweis: Zunächst gilt $\mathfrak{M} = \mathfrak{M}\varepsilon_k$, also $\mathfrak{M} = \mathfrak{M}\mathfrak{R}_k + \sum \mathfrak{M}\mathfrak{R}\varepsilon_p$, $p \neq k$. Daß hierbei $\mathfrak{M}$ ein $\mathfrak{R}_k$-Rechtsmodul ist, zeigt man ebenso, wie oben $\mathfrak{R}_i \mathfrak{M} = \mathfrak{M}$ bewiesen wurde, nach Vertauschung der linken Seite mit der rechten. Wäre nun $\mathfrak{M}\mathfrak{R}\varepsilon_p \neq 0$, dann wäre, wegen der Treue von $\mathfrak{m}$, $0 \neq \mathfrak{M}\mathfrak{R}\varepsilon_p \mathfrak{m} = \mathfrak{M}\mathfrak{R}\varepsilon_p \sum \mathfrak{d}_j = \mathfrak{M}\mathfrak{R}\varepsilon_p \mathfrak{d}_u = \mathfrak{M}\mathfrak{R}\varepsilon_p \varepsilon_k \mathfrak{d}_u = 0$, q.e.a. — Wegen $\mathfrak{M}\mathfrak{d}_j = 0$ für $j \neq u$, wegen $\mathfrak{M}\mathfrak{d}_u = \mathfrak{M}\mathfrak{m} = \mathfrak{W}_i^{e_i - 1}\mathfrak{m} = \mathfrak{m}_1$ und wegen der Treue von $\mathfrak{m}$ induzieren die Elemente des Ideales $\mathfrak{M} = \mathfrak{M}_{ik}$ lauter K-lineare Abbildungen von $\mathfrak{m}_u - \mathfrak{m}_{u-1}$ in $\mathfrak{m}_1$ derart, daß nur das Nullelement von $\mathfrak{M}$ die Nullabbildung ergibt. Nach der Voraussetzung von Satz 5 ist der Ring K von $\mathfrak{R}$-Endomorphismen von $\mathfrak{m}$ so groß, daß er insbesondere für $\mathfrak{m}_u - \mathfrak{m}_{u-1}$ und für $\mathfrak{m}_1$ deren volle Zentralisatoren $K^{(k)}$ bzw. $K^{(i)}$ induziert. Hierauf lassen sich nun das Lemma aus JACOBSON [7], II.2 und, etwas modifiziert, die darauf folgenden, bei JACOBSON zum Beweis seines Theorems II.2.1 führenden Überlegungen anwenden. Sie ergeben, daß $\mathfrak{M}$ die dichte Gruppe K-linearer Abbildungen von $\mathfrak{m}_u - \mathfrak{m}_{u-1}$ in $\mathfrak{m}_1$ induziert. Wäre nun $\mathfrak{M}$ nicht minimales Ideal, würde also ein Ideal $0 \neq \mathfrak{M}' < \mathfrak{M}$ existieren, dann ließe sich das obige Resultat von JACOBSON auch auf $\mathfrak{M}'$ anwenden, also wäre $\mathfrak{M}'$ ebenfalls die dichte Gruppe, also wäre doch $\mathfrak{M}' = \mathfrak{M}$. Damit ist Satz 5 bewiesen.

Der Satz 5 und sein Zusatz zeigen, daß bei genügend großem K die minimalen Ideale eines distributiv und K-linear darstellbaren Ringes $\mathfrak{R}$ sehr eng mit der Halbordnungsstruktur der verbindungsirreduziblen $\mathfrak{R}$-Untermoduln eines distributiven $\mathfrak{R}$-K-Moduls $\mathfrak{m}$ zusammenhängen; wenn sich diese minimalen Ideale auch bereits durch den Satz 5 (ohne den Zusatz) innerhalb des Ringes $\mathfrak{R}$ beschreiben lassen.

Für einen primären Ring $\mathfrak{R}$, also für $q = 1$, $\varepsilon_1 = 1$, sind in Satz 5 die Gln. (3) trivial erfüllt, also ist $\mathfrak{M}_{11}$ das einzige minimale Ideal in $\mathfrak{R}$. Da $\mathfrak{W}^{e-1}$ den

Gleichungen $\mathfrak{W}\mathfrak{W}^{e-1}=0=\mathfrak{W}^{e-1}\mathfrak{W}$ genügt und ein Ideal ist, ist $\mathfrak{M}_{11}=\mathfrak{W}^{e-1}$, ein Resultat, das bereits in Teil I, Satz 3, auf anderem Wege hergeleitet wurde.

Da $\mathfrak{M}_{ii}=\mathfrak{W}_i^{q_i-1}$ das minimale Ideal des primären Ringes $\mathfrak{R}_i$ ist, sind die $\mathfrak{M}_{ii}$ stets ungleich 0. Sie brauchen aber keine Ideale, die Gln. (3) also in $\mathfrak{R}$ nicht erfüllt zu sein, wie das folgende Beispiel 1 zeigen wird. Andererseits können für ein Paar (i, k), $i \neq k$, die Gln. (3) gelten, ohne daß $\mathfrak{M}_{ik} \neq 0$ ist, z.B. dann, wenn $\mathfrak{R}$ die direkte Summe der Ringe $\mathfrak{R}_i$ ist.

BEISPIEL 1. Der Ring $\mathfrak{R}$ bestehe aus den Dreiecksmatrizen

$$\tau=\begin{pmatrix} p_1 & r_{12} & r_{13} & p_{14} & p_{15} & r_{16} \\ & q_2 & q_{23} & r_{24} & r_{25} & q_{26} \\ & & q_2 & r_{34} & r_{35} & q_{36} \\ & & & p_1 & p_{45} & r_{46} \\ & 0 & & & p_1 & r_{56} \\ & & & & & q_2 \end{pmatrix},$$

deren Matrixelemente p, q, r einem kommutativen Körper K entnommen seien. In $\mathfrak{R}=\mathfrak{R}_1+\mathfrak{R}_2+\mathfrak{N}$ besteht der primäre Ring $\mathfrak{R}_1$ aus denjenigen τ, für die alle q und r gleich 0 sind, der Ring $\mathfrak{R}_2$ aus denjenigen τ, für die alle p und r gleich 0 sind, während der K-Modul $\mathfrak{N}$ gerade diejenigen Matrizen τ enthält, für die alle p und q gleich 0 sind. Bezeichnet man die bekannten Matrixeinheiten durch C_{ik}, $1 \leqq i, k \leqq 6$, dann sind $\mathfrak{M}_{11}=C_{14}K$ und $\mathfrak{M}_{22}=C_{26}K$ keine Ideale n $\mathfrak{R}$, sondern $\mathfrak{M}_{12}=C_{16}K$ ist das einzige minimale Ideal in $\mathfrak{R}$.

Auf das Beispiel 1 werden wir noch später zurückkommen.

Dem Satz 5 kann man etwas mehr über die idealtheoretische Struktur eines distributiv und K-linear darstellbaren Ringes entnehmen, wenn man beachtet, daß jeder $\mathfrak{R}$-Untermodul u eines distributiven $\mathfrak{R}$-K-Moduls m und auch der dazugehörige Faktormodul $\mathfrak{m}-\mathfrak{u}$ ebenfalls wieder distributive $\mathfrak{R}$-Untermodulverbände besitzen und die Faktorringe $\mathfrak{R}/(0:\mathfrak{u})$ bzw. $\mathfrak{R}/(\mathfrak{u}:\mathfrak{m})$ treu darstellen. Wenn z.B. $\mathfrak{R}$ primär ist und $\mathfrak{u}=\mathfrak{W}^{e-k}\mathfrak{m}$ genommen wird, ist nach Teil I, Hilfssatz 3, $(\mathfrak{u}:\mathfrak{m})=(0:\mathfrak{W}^k)=\{\tau \mid \tau \in \mathfrak{R},\ \mathfrak{W}^k\tau=0\}=\mathfrak{A}_k$ innerhalb des Ringes $\mathfrak{R}$ beschreibbar. Der Satz 5 sagt dann etwas über die minimalen Ideale in $\mathfrak{R}/\mathfrak{A}_k$, also über die minimalen Oberideale von $\mathfrak{A}_k$ in $\mathfrak{R}$ aus.

§ 3

Eine rein ringtheoretische Charakterisierung der distributiv und K-linear darstellbaren Ringe läßt sich leicht angeben: Nach Teil I, Satz 2, ist jeder distributive $\mathfrak{R}$-Modul m zyklisch, $\mathfrak{m}=\mathfrak{R}b$, also m als $\mathfrak{R}$-Modul isomorph zu dem mit dem Linksideal $\mathfrak{l}=(0:b)$ gebildeten Faktormodul $\mathfrak{R}-\mathfrak{l}$. Dieser ist dann und nur dann treuer $\mathfrak{R}$-Modul, wenn (0) das einzige in $\mathfrak{l}$ enthaltene Ideal ist. Schreibt man wie in den beiden vorangehenden Paragraphen

$$\mathfrak{R}=\mathfrak{R}_1+\cdots+\mathfrak{R}_q+\mathfrak{N}, \tag{1}$$

so ist nach Teil I, Satz 1, der Verband $V(\mathfrak{m})$ der $\mathfrak{R}$-Untermoduln eines treuen $\mathfrak{R}$-Moduls $\mathfrak{m}$ genau dann distributiv, wenn er die Länge $l=e_1+\cdots+e_q$ hat, dabei ist e_i der Exponent des Radikales $\mathfrak{W}_i$ des primären Ringes $\mathfrak{R}_i=\varepsilon_i\mathfrak{R}\varepsilon_i$. Dies sei notiert als

SATZ 6. *Der in der Form* (1) *zerlegte Artin-Ring mit Einselement ist dann und nur dann distributiv und K-linear darstellbar, wenn* $\mathfrak{R}$ *ein Linksideal* $\mathfrak{l}$ *mit folgenden drei Eigenschaften besitzt: 1. Der* $\mathfrak{R}$*-Linksmodul* $\mathfrak{R}-\mathfrak{l}$ *ist als* $\mathfrak{R}$*-K-Modul erklärbar, 2.* $(\mathfrak{l}:\mathfrak{R})=0$, *das heißt* $\mathfrak{l}$ *enthält nur* (0) *als Ideal, 3. die maximalen Ketten von Linksoberidealen von* $\mathfrak{l}$ *haben die Länge* $l=e_1+\cdots+e_q$.

Setzt man von vornherein voraus, daß $\mathfrak{R}$ nur die in der Behauptung des Satzes 5 genannten $\mathfrak{M}_{ik}\neq 0$ als minimale Ideale besitzt, so läßt sich die Forderung 2. des Satzes 6 dadurch nachprüfen, daß man untersucht, ob eines dieser $\mathfrak{M}_{ik}$ in $\mathfrak{l}$ enthalten ist.

Die zu distributiven $\mathfrak{R}-K$-Moduln $\mathfrak{R}-\mathfrak{l}$ führenden Linksideale $\mathfrak{l}$ sind für primäre Ringe bereits in Teil I, §2 näher beschrieben worden. Die darauf beruhende, verhältnismäßig komplizierte ringinterne Charakterisierung der primären, distributiv darstellbaren Ringe wird im nächsten Paragraphen wesentlich vereinfacht werden. Das analoge Problem für nichtprimäre Artin-Ringe mit Einselement soll noch in diesem Paragraphen darauf zurückgeführt werden. Dabei spielt der aus der Zerlegung (1) des Ringes $\mathfrak{R}$ resultierende Unterring

$$\mathfrak{S}=\mathfrak{R}_1\oplus\cdots\oplus\mathfrak{R}_q \tag{2}$$

eine Rolle. $\mathfrak{S}$ ist direkte Summe der primären Ringe $\mathfrak{R}_i$ im ringtheoretischen Sinne, weil die Idempotente ε_i orthogonal sind. In Teil I, Satz 12 wurde erkannt, daß ein $\mathfrak{R}$-K-Modul $\mathfrak{m}$ bereits dann distributiv ist, wenn er den Ring $\mathfrak{S}$ distributiv darstellt und als $\mathfrak{R}$-Modul treu ist.

$\mathfrak{m}$ sei nun ein distributiver $\mathfrak{R}$-K-Modul. Dann ist $\varepsilon_i\mathfrak{m}$ ein distributiver $\mathfrak{R}_i$-K-Modul, also $\varepsilon_i\mathfrak{m}=\mathfrak{R}_i b^{(i)}$ für ein erzeugendes Element $b^{(i)}$ des $\mathfrak{R}_i$-Moduls $\varepsilon_i\mathfrak{m}$, $i=1,\ldots,q$. Das Element $b=b^{(1)}+\cdots+b^{(q)}$ erzeugt den $\mathfrak{R}$-Modul $\mathfrak{R}b=\mathfrak{m}$, da $\mathfrak{R}b\geqq\sum\mathfrak{R}_i b^{(i)}=\sum\varepsilon_i\mathfrak{m}=\mathfrak{m}$. Demnach gelten

$$\mathfrak{m}\cong\mathfrak{R}-\mathfrak{l},\quad \mathfrak{l}=(0:b)\quad\text{als }\mathfrak{R}\text{-Modul} \tag{3}$$

und

$$\varepsilon_i\mathfrak{m}\cong\mathfrak{R}_i-\mathfrak{l}_i,\quad \mathfrak{l}_i=\{\tau\,|\,\tau\in\mathfrak{R}_i,\ \tau b^{(i)}=0\}\quad\text{als }\mathfrak{R}_i\text{-Moduln.} \tag{4}$$

Wegen $b^{(i)}=\varepsilon_i b$ ist

$$\mathfrak{l}=\{\tau\,|\,\tau\in\mathfrak{R}_i,\ \tau b=0\}=\mathfrak{l}\cap\mathfrak{R}_i. \tag{5}$$

Das in (3) definierte Linksideal $\mathfrak{l}$ hat also die Eigenschaft, daß für $i=1,\ldots,q$ die $\mathfrak{R}_i$-Moduln $\mathfrak{R}_i-(\mathfrak{l}\cap\mathfrak{R}_i)$ die Ringe $\mathfrak{R}_i$ distributiv darstellen. Durch Linksmultiplikation von (3) mit ε_i folgt $\varepsilon_i\mathfrak{m}\cong\varepsilon_i\mathfrak{R}-\varepsilon_i\mathfrak{l}$ als $\mathfrak{R}_i$-Modul. Dies gibt,

zusammen mit (4) einen $\mathfrak{R}_i$-Isomorphismus

$$H_i: \quad \mathfrak{R}_i - \mathfrak{l}_i \cong \varepsilon_i \mathfrak{R} - \varepsilon_i \mathfrak{l}, \tag{6}$$

der wegen $\mathfrak{R}_i \subseteqq \varepsilon_i \mathfrak{R}$ aus der Abbildung

$$H_i: \quad \varrho_i \bmod \mathfrak{l}_i \to \varrho_i \bmod \mathfrak{l} \tag{7}$$

besteht. Der Isomorphismus (6) führt auf $\varepsilon_i \mathfrak{R} = \mathfrak{R}_i \cup \varepsilon_i \mathfrak{l}$, also wegen $\varepsilon_1 + \cdots + \varepsilon_q = 1$ auf

$$\mathfrak{R} = \mathfrak{S} \cup \mathfrak{l}. \tag{8}$$

Mit $\mathfrak{m}$ sind auch $\varepsilon_i \mathfrak{m}$ und $\mathfrak{R}_i - \mathfrak{l}_i$ Doppelmoduln mit K als einem Rechtsoperatorenbereich.

Die Umkehrung dieser Überlegungen wird den Beweis des folgenden Satzes vollenden.

SATZ 7. *Voraussetzungen.* $\mathfrak{R}$ *sei ein Artin-Ring mit Einselement und*

$$\mathfrak{R} = \mathfrak{R}_1 + \cdots + \mathfrak{R}_q + \mathfrak{N}, \tag{1}$$

$$\mathfrak{S} = \mathfrak{R}_1 \oplus \cdots \oplus \mathfrak{R}_q. \tag{2}$$

Behauptung. $\mathfrak{R}$ *ist dann und nur dann distributiv und K-linear darstellbar, wenn* $\mathfrak{R}$ *ein Linksideal* $\mathfrak{l}$ *mit folgenden Eigenschaften besitzt: 1. Der* $\mathfrak{R}$*-Modul* $\mathfrak{R} - \mathfrak{l}$ *ist als* $\mathfrak{R}$*-K-Modul erklärbar. 2.* $\mathfrak{l}$ *enthält nur* 0 *als Ideal. 3.* $\mathfrak{R} = \mathfrak{S} \cup \mathfrak{l}$. *4.* $\mathfrak{R}_i - (\mathfrak{l} \cap \mathfrak{R}_i)$ *stellt den primären Ring* $\mathfrak{R}_i$ *distributiv und K-linear dar,* $i = 1, \ldots; q$. *Falls das Linksideal* $\mathfrak{l}$ *die Forderungen 1. bis 4. erfüllt, ist* $\mathfrak{R} - \mathfrak{l}$ *ein distributiver* $\mathfrak{R}$*-K-Modul.*

BEWEIS. Die Bedingungen 1. bis 4. seien für $\mathfrak{l}$ erfüllt. Dann ist $\mathfrak{R} - \mathfrak{l}$ nach 1. $\mathfrak{R}$-K-Modul, der wegen 2. treu bezüglich $\mathfrak{R}$ ist. Zu zeigen bleibt nach Teil I, Sätze 11 und 12, daß die Moduln $\varepsilon_i(\mathfrak{R} - \mathfrak{l})$ die Ringe $\mathfrak{R}_i$ distributiv darstellen. Nach 4. haben die $\mathfrak{R}_i$-Moduln $\mathfrak{R}_i - (\mathfrak{l} \cap \mathfrak{R}_i)$ diese Eigenschaft. Wegen $\mathfrak{R}_i = \varepsilon_i \mathfrak{R} \varepsilon_i \subseteqq \varepsilon_i \mathfrak{R}$ ist die Abbildung

$$H_i: \quad \varrho_i \bmod (\mathfrak{l} \cap \mathfrak{R}_i) \to \varrho_i \bmod \varepsilon_i \mathfrak{l}$$

ein $\mathfrak{R}_i$-Monomorphismus von $\mathfrak{R}_i - \mathfrak{l}_i$ in $\varepsilon_i \mathfrak{R} - \varepsilon_i \mathfrak{l}$. Diese Abbildung ist ein Epimorphismus, wenn $\mathfrak{R}_i \cup \varepsilon_i \mathfrak{l} = \varepsilon_i \mathfrak{R}$ ist, eine Forderung, die nach 3. wegen $\mathfrak{R}_i = \varepsilon_i \mathfrak{S}$ erfüllt ist. Q.e.d.

Die die Treue von $\mathfrak{R} - \mathfrak{l}$ betreffende Eigenschaft 2. läßt sich wieder mittels des Satzes 5 nachprüfen, falls man für $\mathfrak{R}$ voraussetzt, daß die $\mathfrak{M}_{ik} \neq 0$ aus Satz 5 gerade die minimalen Ideale von $\mathfrak{R}$ sind. Bereits aus der Forderung 4. folgt, daß $\mathfrak{R} - \mathfrak{l}$ die $\mathfrak{M}_{ii}$ treu darstellt, denn die $\mathfrak{M}_{ii}$ sind nach §2 die minimalen Ideale $\mathfrak{P}_i^{e_i - 1}$ der primären Ringe $\mathfrak{R}_i$. Trotzdem ist 2. nicht überflüssig. Auch dies zeigt das

BEISPIEL 1 aus §2. Die diesen Ring aus $\mathfrak{R}$ definierenden Matrizen beziehen sich auf einen distributiven $\mathfrak{R}$-K-Modul $\mathfrak{m}$ mit Basiselementen $b_1, \ldots, b_6$. Er ist durch $b = b_5 + b_6$ erzeugbar, $\mathfrak{m} = \mathfrak{R} b$. Das Linksideal $\mathfrak{l} = (0 : b)$ besteht

aus den Matrizen

$$\lambda = \begin{pmatrix} p_1 & r_{12} & r_{13} & p_{14} & p_{15}, & -p_{15} \\ & q_2 & q_{23} & r_{24} & r_{25}, & -r_{25} \\ & & q_2 & r_{34} & r_{35}, & -r_{35} \\ & & & p_1 & p_{45}, & -p_{45} \\ & & & & p_1, & -p_1 \\ & & & & & 0 \end{pmatrix}$$

und erfüllt die Forderungen 1. bis 4. von Satz 7. Das Ideal $\mathfrak{l}' = \mathfrak{l} + C_{16}K$ erfüllt ebenfalls 1. und 3. Ferner ist $\mathfrak{l}' \cap \mathfrak{R}_1 = C_{14}K = \mathfrak{l} \cap \mathfrak{R}_1$ und $\mathfrak{l}' \cap \mathfrak{R}_2 = C_{23}K = \mathfrak{l} \cap \mathfrak{R}_2$, also gilt auch 4. für $\mathfrak{l}'$. Dagegen widerspricht $C_{16}K \subseteqq \mathfrak{l}'$ der Forderung 2.; $\mathfrak{R} - \mathfrak{l}'$ ist also nicht treu.

Schließlich sei noch bemerkt, daß zwar, wie oben gezeigt, jeder distributive $\mathfrak{R}$-K-Modul $\mathfrak{m}$ isomorph ist zu $\mathfrak{R} - \mathfrak{l}$ mit einem die Eigenschaften 1. bis 4. von Satz 7 besitzenden Linksideal $\mathfrak{l}''$, daß aber nicht jedes Linksideal $\mathfrak{l}''$ mit $\mathfrak{R} - \mathfrak{l}'' \cong \mathfrak{m}$ diese Eigenschaften zu besitzen braucht. Auch das läßt sich an dem obigen Beispiel 1. zeigen: $\mathfrak{m}$ ist offenbar auch $\mathfrak{R}$-isomorph zu dem $\mathfrak{R}$-Modul $\mathfrak{R} - \mathfrak{l}''$, dessen Linksideal $\mathfrak{l}''$ aus denjenigen Matrizen τ besteht, in deren letzter Spalte ausschließlich Nullen auftreten. Dann ist bereits 3. also $\mathfrak{R} = \mathfrak{S} \cup \mathfrak{l}''$ nicht mehr richtig.

§ 4

Der primäre Ring $\mathfrak{R}$ sei cleft, das heißt mit Hilfe des vollen Matrizenringes K_d über dem Körper K sei $\mathfrak{R}$ zerlegbar in

$$\mathfrak{R} = K_d + \mathfrak{W} \tag{1}$$

und dabei e der Exponent des Radikales $\mathfrak{W}$ von $\mathfrak{R}$. In $\overline{\mathfrak{R}} = \mathfrak{R}/\mathfrak{W}$ sei für das Idempotent $\zeta \in \mathfrak{R}$ mit $\zeta K = K\zeta$ das Rechtsideal $\overline{\zeta}\overline{\mathfrak{R}}$ maximal und damit $\mathfrak{r} = \zeta\mathfrak{R} \cup \mathfrak{W}$ ein maximales Rechtsideal in $\mathfrak{R}$. Für $d = 1$, das heißt für einen vollständig primären Ring, ist $\zeta = 0$. Setze $\eta = 1 - \zeta$, also $\eta^2 = \eta \neq 0$. Dann ist $\overline{\mathfrak{R}}\overline{\eta}$ ein minimales Linksideal in $\overline{\mathfrak{R}}$, also $\overline{\eta}\overline{\mathfrak{R}}\overline{\eta} = \overline{\eta}K = K\overline{\eta}$. Durch Auflösung der Restklassen modulo $\mathfrak{W}$ erhält man

$$\eta\,\mathfrak{R}\,\eta \cup \mathfrak{W} = \eta K \cup \mathfrak{W} = K\eta \cup \mathfrak{W}. \tag{2}$$

Nach §2 ist $\mathfrak{W}^{e-1}$ das einzige minimale Ideal in $\mathfrak{R}$, falls $\mathfrak{R}$ distributiv und K-linear darstellbar ist. Wie in Behrens [4], Sätze 3 und 4 gezeigt, ist dann $\mathfrak{W}^{e-1}\eta \neq 0$, und $\mathfrak{R}$ wird durch einen Faktormodul $\mathfrak{R} - \mathfrak{l}$ distributiv dargestellt, dessen Linksideal $\mathfrak{l}$ ein $\mathfrak{l}K \subseteqq \mathfrak{l}$ genügendes Oberideal von $\mathfrak{R}\zeta$ mit folgender Maximaleigenschaft ist: $\mathfrak{l}$ genügt der Gleichung

$$\mathfrak{l} \cap (0 : \mathfrak{r}) = 0, \tag{3}$$

aber jedes echte Oberideal von $\mathfrak{l}$ schneidet das Linksideal $(0:\mathfrak{r})$ nicht nur in 0. Nach Satz 1 folgt aus $\mathfrak{l}K \subseteqq \mathfrak{l}$, daß auch $\mathfrak{l}'K \subseteqq \mathfrak{l}'$ für jedes Linksoberideal $\mathfrak{l}'$ von $\mathfrak{l}$ richtig ist.

Diese Eigenschaften von $\mathfrak{l}$ genügen nun aber, wie ebenfalls in Teil I, §2, durch das Beispiel 2 bewiesen wurde, noch nicht, um diejenigen Linksideale $\mathfrak{l}$ zu charakterisieren, die zu distributiven $\mathfrak{R}$-K-Moduln $\mathfrak{R}-\mathfrak{l}$ führen. Alle weiteren Untersuchungen in §2 jener Arbeit laufen darauf hinaus, die obigen Eigenschaften zu den genauen an $\mathfrak{l}$ zu richtenden Forderungen zu verschärfen. Die Charakterisierung von $\mathfrak{l}$ läßt sich vereinfachen, wenn man den letzten Absatz von §2 der vorliegenden Abhandlung heranzieht. Danach stellt nämlich $\mathfrak{m}-\mathfrak{W}^{e-p}\mathfrak{m}$ den Faktorring $\mathfrak{R}/(0\cdot\mathfrak{W}^p)$ distributiv und K-linear dar, wenn $\mathfrak{m}$ ein distributiver $\mathfrak{R}$-K-Modul ist. Der Quotient $(0\cdot\mathfrak{W}^p)=\{\mathfrak{r}\,|\,\mathfrak{r}\in\mathfrak{R},\ \mathfrak{W}^p\mathfrak{r}=0\}$ sei gleich A_p gesetzt, also $A_0=0$, $A_{e-1}=\mathfrak{W}$, $A_e=\mathfrak{R}$. Wegen $\mathfrak{W}^{e-1}\neq 0$ ist dabei der Exponent des Radikales $\mathfrak{W}$ mod $\mathfrak{A}_p$ von $\mathfrak{R}/\mathfrak{A}_p$ gleich $e-p$. Auch für die Ringe $\mathfrak{R}/\mathfrak{A}_p$ gelten daher die obigen Bemerkungen, nämlich

1. $\mathfrak{W}^{e-p-1}$ mod $\mathfrak{A}_p$ ist das einzige minimale Ideal in $\mathfrak{R}/\mathfrak{A}_p$.

2. $\mathfrak{W}^{e-p-1}\eta\subsetneqq\mathfrak{A}_p$.

3. In $\mathfrak{R}$ existiert ein Linksideal $\mathfrak{l}$ mit den folgenden Eigenschaften: $\alpha)$ $\mathfrak{l}\supseteq\mathfrak{R}\zeta$, $\beta)$ $\mathfrak{l}'K\subseteqq\mathfrak{l}'$ für jedes Linksoberideal $\mathfrak{l}'$ von $\mathfrak{l}'$, $\gamma)$ für alle $p=0,1,\ldots,e-1$ ist $\mathfrak{l}\cup\mathfrak{A}_p$ bezüglich der Forderung

$$(\mathfrak{l}\cup\mathfrak{A}_p)\cap(\mathfrak{A}_p:\mathfrak{r})=\mathfrak{A}_p \tag{4}$$

unter allen Linksoberidealen von $\mathfrak{l}\cup\mathfrak{A}_p$ maximal.

Offenbar ist (4) die Gl. (3) modulo $\mathfrak{A}_p$ betrachtet, das Ideal $\mathfrak{R}\zeta$ erfüllt wegen $\zeta\in\mathfrak{r}$ die Gl. (4) für $p=0$. In Hilfssatz 2 wird gezeigt werden, daß ein Linksideal $\mathfrak{l}$, welches der Gl. (4) für $p=0$ genügt, diese Gleichung für alle p befriedigt. Selbstverständlich gibt es für jeden primären Ring — auch für die nicht distributiv darstellbaren — und jedes p ein von p abhängendes Linksoderideal $\mathfrak{l}_p$ von $\mathfrak{R}\zeta$, welches bezüglich der Forderung (4) unter allen Linksoberidealen von $\mathfrak{A}_p$ maximal ist. In $3\,\gamma)$ wird jedoch verlangt, daß $\mathfrak{l}_p=\mathfrak{l}\cup\mathfrak{A}_p$ mit einem einzigen $\mathfrak{l}$ für alle p richtig ist. Neben 1. und 2. ist die letzte zur Charakterisierung der distributiv darstellbaren unter den primären Ringen benötigte Eigenschaft gerade die Existenz eines 3. befriedigenden $\mathfrak{l}$, wie die folgenden Ausführungen zeigen werden.

HILFSSATZ 2. *Voraussetzungen.* $\mathfrak{R}$ *sei ein primärer Artin-Ring mit Einselement,* $\mathfrak{A}_p=(0\cdot\mathfrak{W}^p)$, $\mathfrak{r}$ *ein Rechtsideal und* $\mathfrak{l}$ *ein Linksideal in* $\mathfrak{R}$ *mit*

$$\mathfrak{l}\cap(0:\mathfrak{r})=0. \tag{5}$$

Behauptung. $(\mathfrak{l}\cup\mathfrak{A}_p)\cap(\mathfrak{A}_p:\mathfrak{r})=\mathfrak{A}_p$ *für* $p=0,1,\ldots,e$.

BEWEIS. Angenommen, es existierte ein nicht in $\mathfrak{A}_p$ enthaltenes Ringelement σ aus dem Durchschnitt $(\mathfrak{l}\cup\mathfrak{A}_p)\cap(\mathfrak{A}_p:\mathfrak{r})$. Da $\sigma\notin\mathfrak{A}_p$, ist $\mathfrak{W}^p\sigma\neq 0$. Aus $\sigma\in(\mathfrak{A}_p:\mathfrak{r})$ folgt $\sigma\mathfrak{r}\subseteqq\mathfrak{A}_p$, also $\mathfrak{W}^p\cdot\sigma\mathfrak{r}=0$, also auch $\mathfrak{W}^p\sigma\cdot\mathfrak{r}=0$. Schließlich zieht $\sigma\in\mathfrak{l}\cup\mathfrak{A}_p$ nach sich $\mathfrak{W}^p\sigma\subseteqq\mathfrak{W}^p\mathfrak{l}\cup 0=\mathfrak{W}^p\mathfrak{l}$. Dies widerspricht aber wegen $\mathfrak{W}^p\sigma\neq 0$ der Gl. (5).

Der primäre Ring $\mathfrak{R}$ genüge im folgenden den Forderungen 1. und 2. In ihm sei $\mathfrak{l}$ ein Linksideal mit den unter 3. genannten Eigenschaften. Die Gl. (19) in Teil I legt nahe den

HILFSSATZ 3.

(6) $$\mathfrak{W}^{e-p}\eta \cup \mathfrak{l} = \mathfrak{A}_p \cup \mathfrak{l} \quad \textit{für alle } p.$$

Beweis durch vollständige Induktion nach p. Wegen $\mathfrak{W}^e = 0$ liegt für alle p die linke Seite von (6) in der rechten. Die Verankerung der Induktion bei $p = 0$ ist wegen $\mathfrak{A}_0 = 0$ trivial. Also gelte (6) für p. Angenommen, es wäre $\mathfrak{A}_{p+1} \subseteqq \mathfrak{W}^{e-p-1}\eta \cup \mathfrak{l}$ oder, äquivalent damit, es wäre

(7) $$\mathfrak{A}_{p+1} > (\mathfrak{W}^{e-p-1}\eta \cup \mathfrak{l}) \cap \mathfrak{A}_{p+1} = \mathfrak{W}^{e-p-1}\eta \cup (\mathfrak{l} \cap \mathfrak{A}_{p+1}),$$

wobei die letzte Gleichung aus der Modularität des Verbandes der Linksideale eines Ringes und aus $\mathfrak{W}^{e-p-1}\eta \subseteqq \mathfrak{A}_{p+1}$ folgt. Nach Induktionsannahme enthält die rechte Seite von (7) das Ideal $\mathfrak{A}_p$. Weil $\mathfrak{r} = \zeta\mathfrak{R} \cup \mathfrak{W}$ und $\eta\zeta = 0$, liegt $\mathfrak{W}^{e-p-1}\eta$ in $(\mathfrak{A}_p : \mathfrak{r})$. Demnach ist $\mathfrak{b} = \mathfrak{W}^{e-p-1}\eta \cap (\mathfrak{l} \cap \mathfrak{A}_{p+1}) \subseteqq \mathfrak{A}_p$, denn nach Voraussetzung gilt (4). Modulo $\mathfrak{A}_p$ betrachtet, ist also die letzte Verbindung in (7) eine $\mathfrak{R}$-direkte Summe von Linksidealen. Schließlich ist auch $\mathfrak{W}\mathfrak{A}_{p+1} \subseteqq \mathfrak{A}_p$. Aus der vollen Reduzibilität des $\mathfrak{R}/\mathfrak{W}$-Moduls $\mathfrak{A}_{p+1} - \mathfrak{A}_p$ folgt die Existenz eines echten Linksoberideales $\mathfrak{b}$ von $\mathfrak{A}_p$ derart, daß

(8) $$\mathfrak{A}_{p+1} = \mathfrak{W}^{e-p-1}\eta + (\mathfrak{l} \cap \mathfrak{A}_{p+1}) + \mathfrak{b} \quad \text{modulo } \mathfrak{A}_p$$

eine $\mathfrak{R}$-direkte Summe von Linksidealen in $\mathfrak{R}/\mathfrak{A}_p$ ist. Nun ist $\mathfrak{b} \subsetneqq \mathfrak{l} \cup \mathfrak{A}_p$, weil sonst wegen der Induktionsannahme (6) für p das Ideal $\mathfrak{A}_{p+1}$ doch in $\mathfrak{W}^{e-p-1}\eta \cup \mathfrak{l}$ enthalten wäre. Für das echte Linksoberideal $\mathfrak{b} \cup \mathfrak{l} \cup \mathfrak{A}_p = \mathfrak{b} \cup \mathfrak{l}$ von $\mathfrak{l} \cup \mathfrak{A}_p$ ist nach Voraussetzung 3 γ) über $\mathfrak{l}$ die Gl. (4) nicht mehr richtig, also $\mathfrak{b}' = (\mathfrak{b} \cup \mathfrak{l}) \cap (\mathfrak{A}_p : \mathfrak{r}) > \mathfrak{A}_p$. Außerdem ist (alles wegen der zu einem Widerspruch zu führenden Annahme, daß die Gl. (6) nicht mehr für $p+1$ gelte) wegen 3 β) die Inklusion $(\mathfrak{b} \cup \mathfrak{l})K \subseteqq \mathfrak{b} \cup \mathfrak{l}$, also auch $\mathfrak{b}'K \subseteqq \mathfrak{b}'$ richtig. Nach den Voraussetzungen 1. und 2. über $\mathfrak{R}$ enthält das echte Oberideal $\mathfrak{b}'\mathfrak{R}$ von $\mathfrak{A}_p$ das Ideal $\mathfrak{W}^{e-p-1}$, also auch $\mathfrak{W}^{e-p-1}\eta$. Daraus und aus (2) folgt $\mathfrak{W}^{e-p-1}\eta \subseteqq \mathfrak{b}'\mathfrak{R}\eta \subseteqq \mathfrak{b}'\zeta\mathfrak{R}\eta \cup \mathfrak{b}'\eta\mathfrak{R}\eta \subseteqq \mathfrak{b}'\mathfrak{W} \cup \mathfrak{b}'(K\eta \cup \mathfrak{W}) \subseteqq \mathfrak{A}_p \cup \mathfrak{b}'K \subseteqq \mathfrak{A}_p \cup \mathfrak{b}'$. Dies führt aber zu einem Widerspruch, denn wegen $\mathfrak{A}_p \cup \mathfrak{b}' \subseteqq \mathfrak{b} \cup \mathfrak{l}$ folgt hieraus $\mathfrak{W}^{e-p-1}\eta \subseteqq (\mathfrak{b} \cup \mathfrak{l}) \cap \mathfrak{A}_{p+1} = \mathfrak{b} \cup (\mathfrak{l} \cap \mathfrak{A}_{p+1})$, also wäre modulo $\mathfrak{A}_p$ die Summe auf der rechten Seite von (8) doch nicht direkt, da nach Voraussetzung $\mathfrak{W}^{e-p-1}\eta \subsetneqq \mathfrak{A}_p$, q.e.a.

Aus dem Hilfssatz 3 kann man schließen, daß $\mathfrak{R} - (\mathfrak{W}^{e-p} \cup \mathfrak{l})$ *ein treuer* $\mathfrak{R}/\mathfrak{A}_p$*-Modul ist*, das heißt, daß dieser Modul nicht das minimale Ideal $\mathfrak{W}^{e-p-1}$ modulo $\mathfrak{A}_p$ annulliert. $\mathfrak{W}^{e-p-1}\eta$ liegt nämlich wegen $\mathfrak{r} = \zeta\mathfrak{R} \cup \mathfrak{W}$ in $(\mathfrak{A}_p : \mathfrak{r})$, jedoch nach Voraussetzung 2. nicht in $\mathfrak{A}_p$. Also ist, wegen (4), $\mathfrak{W}^{e-p-1}\eta \subsetneqq \mathfrak{l} \cup \mathfrak{A}_p$, also gilt erst recht $\mathfrak{W}^{e-p-1} = \mathfrak{W}^{e-p-1} \subsetneqq \mathfrak{l} \cup \mathfrak{A}_p = \mathfrak{W}^{e-p}\eta \cup \mathfrak{l} = \mathfrak{W}^{e-p} \cup \mathfrak{l}$, da $\mathfrak{R}\zeta \subseteqq \mathfrak{l}$. Das $\mathfrak{l}$ enthaltende Linksideal $\mathfrak{W}^{e-p} \cup \mathfrak{l}$ ist nach 3 β) ein K-Rechtsmodul, der Faktormodul $\mathfrak{R} - (\mathfrak{W}^{e-p} \cup \mathfrak{l})$ also ein bezüglich $\mathfrak{R}/\mathfrak{A}_p$, treuer $\mathfrak{R}/\mathfrak{A}_p$-$K$-Modul.

Jedoch sind die echten $\mathfrak{R}/\mathfrak{A}_p$*-K-Faktormoduln von* $\mathfrak{R} - (\mathfrak{W}^{e-p} \cup \mathfrak{l})$ *nicht mehr treu.*

BEWEIS. $\mathfrak{l}'$ sei echtes Linksoberideal von $\mathfrak{W}^{e-p} \cup \mathfrak{l} = \mathfrak{A}_p \cup \mathfrak{l}$, also $\mathfrak{l}'K \subseteqq \mathfrak{l}'$. Nach Voraussetzung 3 γ) über die Maximalität von $\mathfrak{l} \cup \mathfrak{A}_p$ ist $\mathfrak{b}'' = \mathfrak{l}' \cap (\mathfrak{A}_p : \mathfrak{r}) > \mathfrak{A}_p$. Ähnlich wie beim Beweis des Hilfssatzes 3 kann man aus $\mathfrak{b}'' \subseteqq (\mathfrak{A}_p : \mathfrak{r})$ und aus $\zeta \in \mathfrak{r}$ auf $\mathfrak{b}'' \subseteqq \mathfrak{b}''\eta \cup \mathfrak{A}_p$ schließen und dann weiter wegen $\mathfrak{W} \subseteqq \mathfrak{r}$ und (2) auf

$\mathfrak{d}''\,\mathfrak{R} \subseteqq \mathfrak{d}''\,\eta\,\mathfrak{R} \cup \mathfrak{A}_p = \mathfrak{d}''\,\eta\,\mathfrak{R}\zeta \cup \mathfrak{d}''\,\eta\,\mathfrak{R}\eta \cup \mathfrak{A}_p = \mathfrak{d}''\,K\eta \cup \mathfrak{A}_p = \mathfrak{d}''\,\eta \cup \mathfrak{A}_p = \mathfrak{d}'' \cup \mathfrak{A}_p \subseteqq \mathfrak{l}' \cup \mathfrak{A}_p = \mathfrak{l}'$. Demnach wird das in $\mathfrak{R}/\mathfrak{A}_p$ liegende, vom Nullideal verschiedene Ideal $\mathfrak{d}''\,\mathfrak{R}/\mathfrak{A}_p$ durch $\mathfrak{R} - \mathfrak{l}'$ annulliert.

Aus den Ergebnissen der beiden letzten Absätze läßt sich nun die abschließende Behauptung herleiten, daß $\mathfrak{R} - \mathfrak{l}$ ein distributiver $\mathfrak{R}$-K-Modul ist. $\mathfrak{l}$ und damit auch $\mathfrak{R} - \mathfrak{l}$ sind nämlich nach 3 β) $\mathfrak{R}$-K-Moduln. Die Treue von $\mathfrak{R} - \mathfrak{l}$ als $\mathfrak{R}$-Modul ist das auf $p = 0$ spezialisierte Ergebnis des vorletzten Absatzes. Zu zeigen ist gemäß Satz 6 nur noch, daß eine maximale Kette von Linksoberidealen von $\mathfrak{l}$ die Länge e hat oder, was damit äquivalent ist, daß für alle p die Faktormoduln $(\mathfrak{W}^{e-p-1} \cup \mathfrak{l}) - (\mathfrak{W}^{e-p} \cup \mathfrak{l})$ $\mathfrak{R}$-irreduzibel sind. Angenommen, es existierte ein Linksideal $\mathfrak{u}$ mit

$$\mathfrak{W}^{e-p} \cup \mathfrak{l} < \mathfrak{u} < \mathfrak{W}^{e-p-1} \cup \mathfrak{l}. \tag{9}$$

Als Oberideal von $\mathfrak{l}$ wäre $\mathfrak{u}K \subseteqq \mathfrak{u}$; also $\mathfrak{R} - \mathfrak{u}$ ein echter $\mathfrak{R}/\mathfrak{A}_p$-$K$-Faktormodul von $\mathfrak{R} - (\mathfrak{W}^{e-p} \cup \mathfrak{l})$, der nach dem letzten Absatz nicht treu sein könnte, das heißt es wäre $(\mathfrak{u} : \mathfrak{R}) > \mathfrak{A}_p$. Da $(\mathfrak{u} : \mathfrak{R})$ Ideal in $\mathfrak{R}$ ist, folgte aus der Voraussetzung 1 die Inklusion $\mathfrak{W}^{e-p-1} \subseteqq (\mathfrak{u} : \mathfrak{R})$, also $\mathfrak{W}^{e-p-1} \subseteqq \mathfrak{u}$ und damit, wegen $\mathfrak{l} \subseteqq \mathfrak{u}$, die (9) widersprechende Beziehung $\mathfrak{W}^{e-p-1} \cup \mathfrak{l} \subseteqq \mathfrak{u}$.

Alle diese Überlegungen lassen sich zu dem folgenden Satz zusammenfassen:

SATZ 8. *Voraussetzungen* $\mathfrak{R} = K_d + \mathfrak{W}$ *sei ein primärer Artin-Ring. Der Exponent seines Radikales* $\mathfrak{W}$ *sei* e. *Für* $p = 0, 1, \ldots, e-1$ *seien mit* $\mathfrak{A}_p$ *die Ideale* $(0^\cdot . \mathfrak{W}^p) = \{\tau \mid \tau \in \mathfrak{R}, \mathfrak{W}^p \tau = 0\}$ *bezeichnet. — Für das die Gleichung* $\zeta K = K\zeta$ *erfüllende Idempotent* ζ *sei das Rechtsideal* $\mathfrak{r} = \zeta\,\mathfrak{R} \cup \mathfrak{W}$ *maximal in* $\mathfrak{R}$. *Sei* $\eta = 1 - \zeta$.

Behauptungen. Der Ring $\mathfrak{R}$ *ist dann und nur dann distributiv und K-linear darstellbar, wenn folgende drei Bedingungen in ihm für alle* p *erfüllt sind*:

1. $\mathfrak{W}^{e-p-1} \cup \mathfrak{A}_p$ *ist für jedes* p *das einzige minimale Oberideal von* $\mathfrak{A}_p$.
2. $\mathfrak{W}^{e-p-1}\eta \nsubseteqq \mathfrak{A}_p$.
3. *In* $\mathfrak{R}$ *existiert ein Linksideal* $\mathfrak{l}$ *mit folgenden Eigenschaften*
   - α) $\mathfrak{R}\zeta \subseteqq \mathfrak{l}$.
   - β) $\mathfrak{l}'K \subseteqq \mathfrak{l}'$ *für jedes Linksoberideal* $\mathfrak{l}'$ *von* $\mathfrak{l}$.
   - γ) *Für alle* p *gilt mit diesem* $\mathfrak{l}$ *die Gleichung*

$$(\mathfrak{l} \cup \mathfrak{A}_p) \cap (\mathfrak{A}_p : \mathfrak{r}) = \mathfrak{A}_p, \tag{4}$$

aber für kein p *existiert ein echtes Linksoberideal von* $\mathfrak{l} \cup \mathfrak{A}_p$, *welches ebenfalls noch der Gl.* (4) *genügt.*

Wenn die Bedingungen 1. und 2. in $\mathfrak{R}$ *erfüllt sind und für das Linksideal* $\mathfrak{l}$ *auch noch 3. gilt, dann ist* $\mathfrak{R} - \mathfrak{l}$ *ein distributiver* $\mathfrak{R}$-K-*Modul.*

Wenn $\mathfrak{R} - \mathfrak{l}^*$ ein distributiver $\mathfrak{R}$-K-Modul ist, besitzt $\mathfrak{l}^*$ die unter 3. genannten Eigenschaften für ein geeignetes Idempotent ζ.

In jedem primären Artin-Ring mit Einselement impliziert für beliebige Rechts- und Linksideale $\mathfrak{r}$ bzw. $\mathfrak{l}$ die Gl. (4) für $p=0$ ihre Gültigkeit für alle p.

Benutzt man den Begriff des *Sockels eines Ringes* — s. z.B. JACOBSON [7], IV.3 — so kann man die $\mathfrak{A}_p$ und das Linksideal $(0:\mathfrak{r})$ folgendermaßen beschreiben: Der Linkssockel eines Ringes $\mathfrak{R}$ besteht aus der Verbindung der minimalen Linksideale von $\mathfrak{R}$, analog der Rechtssockel. Wegen der Nilpotenz $\mathfrak{W}$ gilt für jedes minimale Linksideal $\mathfrak{l}_1$ von $\mathfrak{R}$ die Gleichung $\mathfrak{W}\mathfrak{l}_1=0$, also $\mathfrak{l}\subseteqq\mathfrak{A}_1$. Als $\mathfrak{R}/\mathfrak{W}$-Modul ist $\mathfrak{A}_1$ vollreduzibel, also direkte Summe minimaler Linksideale von $\mathfrak{R}$. Hieran sieht man, daß $\mathfrak{A}_1=(0^{\cdot}.\mathfrak{W})$ der Linkssockel von $\mathfrak{R}$, und ebenso $\mathfrak{A}_2/\mathfrak{A}_1$ der Linkssockel des Faktorringes $\mathfrak{R}/\mathfrak{A}_1$ ist, und so weiter. — Für das maximale Rechtsideal $\mathfrak{r}=\zeta\mathfrak{R}\cup\mathfrak{W}$ gilt $(0:\mathfrak{r})=(0:\mathfrak{W})\cap(0:\zeta)$. Hierin ist $(0:\mathfrak{W})$ der Rechtssockel des Ringes $\mathfrak{R}$; dieser also gleich $(0:\mathfrak{r})$ falls $\mathfrak{R}$ vollständig primär ist. Wenn $\mathfrak{R}$ nur primär ist, ist wegen $\zeta^2=\zeta$ und $\eta=1-\zeta$ der Quotient $(0:\zeta)=\mathfrak{R}\eta$. Im Faktorring $\overline{\mathfrak{R}}=\mathfrak{R}/\mathfrak{W}$ ist dabei $\overline{\mathfrak{R}}\,\overline{\eta}$ ein minimales Linksideal, und damit $\mathfrak{R}\eta$ selbst ein unzerlegbares Ideal in $\mathfrak{R}$. Hieraus folgt, daß sich das Linksideal $(0:\mathfrak{r})=(0:\mathfrak{W})\cap\mathfrak{R}\eta$ auch als Schnitt des Rechtssockels $(0:\mathfrak{W})$ von $\mathfrak{R}$ mit einem unzerlegbaren Linksideal $\mathfrak{R}\eta$ beschreiben läßt.

Der Satz 8 gibt eine Charakterisierung der distributiv und K-linear darstellbaren primären Ringe $\mathfrak{R}$ durch innere Eigenschaften von $\mathfrak{R}$, die neben die in Teil I am Ende von §2 gegebene ringinterne Beschreibung dieser Ringe tritt. Sie ist vielleicht befriedigender als die Charakterisierung von Teil I, weil in ihr von $\mathfrak{l}$ nur als einem Linksideal gesprochen wird, aber nicht auch noch als einem ω-Modul bezüglich des in Teil I definierten speziellen Ringelementes ω, dessen Existenz dort bei der Charakterisierung postuliert werden mußte. Nur scheinbar vermeiden die Ausführungen in Teil I den schrittweisen Aufbau, der in den Satz 8 der vorliegenden Arbeit dadurch hineinkommt, daß 1. bis 3. für alle p verlangt wird. Man darf nämlich nicht verkennen, daß obwohl 1., 2. und 3 γ) in Teil I nur für $p=0$ gefordert werden, die Einführung von ω einen solchen Stufenaufbau enthält. — Beide ringinterne Kriterien beruhen letzten Endes auf der in Satz 6 ausgesprochenen grundlegenden Charakterisierung der distributiv und K-linear darstellbaren Ringe, die wegen der dort genannten Gleichung $l=e_1+\cdots+e_q$ prinzipiell einen auf den einzelnen Sockeln beruhenden schrittweisen Aufbau rechtfertigt.

Die Frage liegt nahe, ob man in Satz 8 die Bedingung 3 weglassen kann, ob also bereits 1. und 2. für die distributive und K-lineare Darstellbarkeit von $\mathfrak{R}$ hinreichen. Dem ist jedoch nicht so, wie das folgende Beispiel zeigt:

BEISPIEL 2.

$$\mathfrak{R}=\{\tau\},\qquad \tau=\begin{pmatrix} t_1 & t_2 & t_3 & t_4 & t_5\\ & t_1 & 0 & 0 & t_2\\ & & t_1 & t_2 & t_6\\ 0 & & & t_1 & t_7\\ & & & & t_1 \end{pmatrix},$$

t_i beliebig aus einem kommutativen Körper K. Der Exponent von $\mathfrak{W}$ ist 4. Ferner ist $\zeta=0$, also $\mathfrak{r}=\mathfrak{W}$. Die Linkssockeln sind

$$\mathfrak{A}_1=\{\tau\,|\,0=t_1=t_2=t_6=t_7\} \quad \text{in } \mathfrak{R},$$
$$\mathfrak{A}_2=\{\tau\,|\,0=t_1=t_7\} \quad \text{modulo } \mathfrak{A}_1 \text{ in } \mathfrak{R}/\mathfrak{A}_1,$$
$$\mathfrak{A}_3=\mathfrak{W} \quad \text{modulo } \mathfrak{A}_2 \text{ in } \mathfrak{R}/\mathfrak{A}_2.$$

Die einzigen minimalen Oberideale von $\mathfrak{A}_0=0$, $\mathfrak{A}_1$, $\mathfrak{A}_2$ und $\mathfrak{A}_3$ sind der Reihe nach $\mathfrak{W}^3$, $\mathfrak{W}^2\cup\mathfrak{A}_1$, $\mathfrak{W}\cup\mathfrak{A}_2=\mathfrak{W}$ und $\mathfrak{R}$. Der Ring $\mathfrak{R}$ besitzt nur ein einziges maximales, den Quotienten $(0:\mathfrak{r})=(0:\mathfrak{W})=\{\tau\,|\,0=t_1=t_2=t_3=t_4\}$ nur in 0 schneidendes Linksideal, nämlich $\mathfrak{l}=\{\tau\,|\,0=t_1=t_2=t_3=t_6=t\}=C_{13}K+C_{14}K$. Jedoch ist $\mathfrak{R}$ nicht distributiv darstellbar. Nach dem in Satz 8 ausgesprochenen Kriterium müßte nämlich $\mathfrak{l}\cup\mathfrak{A}_1$ ebenfalls maximal sein unter allen Linksoberidealen $\mathfrak{l}'$ von $\mathfrak{l}\cup\mathfrak{A}_1$ bezüglich $(\mathfrak{l}\cup\mathfrak{A}_1)\cap(\mathfrak{A}_1:\mathfrak{r})=\mathfrak{A}_1$. Dies ist aber nicht der Fall, denn das echte Linksoberideal $\mathfrak{l}'=\{\tau\,|\,0=t_1=t_6=t_7\}$ von $\mathfrak{l}\cup\mathfrak{A}_1$ schneidet ebenfalls $(\mathfrak{A}_1:\mathfrak{r})=(\mathfrak{A}_1:\mathfrak{W})=\{\tau\,|\,0=t_1=t_2\}$ nur in $\mathfrak{A}_1$.

Literatur

[1] ARTIN, E., C. J. NESBITT and R. M. THRALL: Rings with minimum condition. Ann. Arbor 1944.

[2] BEHRENS, E.-A.: Algebren mit vorgegebenem endlichem, distributivem Idealverband. Math. Ann. **133**, 79–90 (1957).

[3] — Eine Charakterisierung der T-Moduln mit distributivem Untermodulverband bei halbprimärem T. Arch. Math. **8**, 265–273 (1957).

[4] — Distributiv darstellbare Ringe. Math. Z. **73**, 409–432 (1960).

[5] BIRKHOFF, G.: Lattice Theory, 2. Aufl. New York 1948.

[6] HERMES, H.: Einführung in die Verbandstheorie. Berlin 1955.

[7] JACOBSON, N.: Structure of rings. Providence 1956.

Frankfurt a. M., Mathematisches Seminar der Universität

(Eingegangen am 26. November 1960)

SABIDUSSI, G.
Math. Zeitschr. 76, 385–401 (1961)

Graph Derivatives[1])

By
GERT SABIDUSSI

1. Introduction

By a *graph* X we mean a set $V(X)$ together with a set $E(X)$ of unordered pairs of *distinct* elements of $V(X)$. We shall use brackets to indicate unordered pairs. The elements of $V(X)$ are called the *vertices* of X, the elements of $E(X)$ the *edges* of X. Two edges $e=[x, y]$ and $e'=[x', y']$ of X will be called *adjacent* if and only if $e \neq e'$ and two of the vertices x, y, x', y' are equal.

With every graph X one can associate a graph ∂X, the *derivative*[2]) of X, by defining

$$V(\partial X) = E(X),$$

$$E(\partial X) = \{[e, e']: e \text{ and } e' \text{ are adjacent edges of } X\}.$$

It has been proved by WHITNEY [4] that if X and Y are non-empty connected graphs with $\partial X \approx \partial Y$[3]) and such that neither X nor Y is a complete 3-graph, then $X \approx Y$. For a restricted class of graphs a proof different from Whitney's was recently given by IZBICKI [2].

The present paper is concerned with the relationships between covers of graphs by non-overlapping complete graphs and a certain class of connected even graphs (called admissible in section 4). Considering a graph X as a symmetric, strictly non-reflexive relation on its set of vertices, we form the relational product X^2 of X by itself, and observe that if X is an admissible even graph, then X^2 consists of exactly two components each of which admits a cover by non-overlapping complete graphs. We are interested in deriving conditions under which these natural covers are (i) maximal with respect to refinement, or (ii) invariant under the action of automorphisms. The results obtained (in particular, theorem (4.8)) enable us to prove Whitney's theorem as a simple corollary. We also show that if X is a connected graph not isomorphic to any of four exceptional graphs, then X and ∂X have isomorphic automorphism groups (section 5). In section 7 we prove Whitney's theorem for graphs with multiple edges. For these graphs, however, the validity of the theorem is somewhat reduced.

---

[1]) Written with the support of the National Science Foundation of the United States of America, Grant No. NSF-G 14084.

[2]) ∂X is the dual of X in the sense of combinatorial topology. For our purposes, however, it seemed justified to apply this term to the graph $\delta_{\mathfrak{N}} X$ (cf. action 3).

[3]) For a definition of $\approx$ (isomorphism) see section 2.

2. Definitions

Let X and Y be graphs. By $X \cup Y$ and $X \cap Y$ we mean the graphs given by $V(X \cup Y) = V(X) \cup V(Y)$, $E(X \cup Y) = E(X) \cup E(Y)$, and $V(X \cap Y) = V(X) \cap V(Y)$, $E(X \cap Y) = E(X) \cap E(Y)$, respectively. Let $x \in V(X)$. Then (x) denotes the graph consisting of the vertex x and no edges. Similarly, if $e \notin E(X)$, (e) denotes the graph formed by e and its two ends. When no confusion is likely we write x instead of (x), and e instead of (e). Also, we shall frequently write $x \in X$ and $e \in X$ for $x \in V(X)$ and $e \in E(X)$, respectively.

Let A be a set. By $|A|$ we denote the cardinal of A. If X is a graph $|V(X)|$ is called the *order* of X, and is denoted by $|X|$.

Given a graph X and a vertex $x \in V(X)$ by $E(X; x)$ and $V(X; x)$ we mean the following sets:

$$E(X; x) = \{e \in E(X) : (e) > (x)\},$$
$$V(X; x) = \{y \in V(X) : [x, y] \in E(X)\}.$$

By $d(X; x)$ we shall denote the *degree* of x in X, i.e., $d(X; x) = |E(X; x)| = |V(X; x)|$.

Let X be a graph, $x, y \in V(X)$. A *path* joining x and y is a sequence $P = (x_0, x_1, \ldots, x_n)$, $n \geqq 0$, where $x_0 = x$, $x_n = y$, and $[x_{i-1}, x_i] \in E(X)$, $i = 1, \ldots, n$. n is called the *length* of P and is denoted by $|P|$. A graph X is *connected* if any two vertices of X are joined by a path.

An *n-circuit* of X ($n \geqq 3$) is a sequence $C = (x_1, \ldots, x_n)$ of distinct vertices of X such that $[x_{i-1}, x_i] \in E(X)$, $i = 2, \ldots, n$, $[x_1, x_n] \in E(X)$. $n = |C|$ is called the *order* of C. By mesh X we shall mean $\min |C|$ extended over all circuits of X. If X contains no circuits we put mesh $X = \infty$.

Let X and Y be graphs. A function $\varphi : V(X) \to V(Y)$ is called a *homomorphism* of X into Y if and only if $[\varphi x, \varphi y] \in E(Y)$ whenever $[x, y] \in E(X)$. We shall always write $\varphi : X \to Y$ instead of $\varphi : V(X) \to V(Y)$. Note that if φ is a homomorphism of X into Y then φ induces a function $\tilde{\varphi} : E(X) \to E(Y)$ by

$$\tilde{\varphi}[x, y] = [\varphi x, \varphi y], \qquad [x, y] \in E(X).$$

A homomorphism of X into Y is a *monomorphism* if it is one-one. A monomorphism $\varphi : X \to Y$ is an *isomorphism* of X *onto* Y if *both* φ *and* $\tilde{\varphi}$ are *onto*. Equivalently, φ is an isomorphism if both φ and φ^{-1} are monomorphisms. When no confusion is likely, we shall write φ for $\tilde{\varphi}$. If there is an isomorphism φ of X onto Y we shall write $X \approx Y$ or $\varphi : X \approx Y$ and say that X and Y are *isomorphic*. An *automorphism* of X is an isomorphism of X onto itself. By $G(X)$ we denote the automorphism group of X.

A graph X is called *complete* if and only if $x, y \in V(X)$, $x \neq y$, implies $[x, y] \in E(X)$. If X is complete and $|X| = n$, X is isomorphic to the *complete n-graph* C_n given by $V(C_n) = N$, where N is the set of all ordinals[4]) whose cardinal is $< n$, and $E(C_n) = \{[\alpha, \beta] : \alpha, \beta \in N, \alpha \neq \beta\}$.

[4]) The special choice of N is made only for purposes of standardization.

3. Covers, quasi-derivatives

(3.1) DEFINITION. Let X be a graph. By a *cover* of X is meant a collection $\mathfrak{A}$ of subgraphs of X such that

(i) $\cup\{A: A\in\mathfrak{A}\}=X$;

(ii) $E(A\cap B)=\square$ (= the empty set) if $A, B\in\mathfrak{A}$ and $A\neq B$ (i.e., distinct members of $\mathfrak{A}$ do not overlap).

It is immediate from (3.1) that for each $e\in E(X)$ there is a unique element of $\mathfrak{A}$ which contains e. We shall denote this element of $\mathfrak{A}$ by A_e.

Let $\mathfrak{A}$ be a cover of X, $\varphi\in G(X)$. By $\varphi\mathfrak{A}$ we denote the set $\{\varphi A: A\in\mathfrak{A}\}$. Clearly $\varphi\mathfrak{A}$ is a cover of X. $\mathfrak{A}$ is called *invariant* if and only if $\varphi\mathfrak{A}=\mathfrak{A}$ for each $\varphi\in G(X)$.

Let $\mathfrak{M}$ be a collection of subgraphs of X with the property that if $\mathfrak{N}$ is a tower[5] in $\mathfrak{M}$, then $\cup\mathfrak{N}\in\mathfrak{M}$. A cover $\mathfrak{A}$ of X is called an *$\mathfrak{M}$-cover* if $\mathfrak{A}<\mathfrak{M}$. In particular, if each $A\in\mathfrak{A}$ is complete, then $\mathfrak{A}$ will be called a *complete* cover. By $\boldsymbol{A}_{\mathfrak{M}}$ we denote the collection of all $\mathfrak{M}$-covers of X.

We partially order $\boldsymbol{A}_{\mathfrak{M}}$ by refinement, i.e., we set $\mathfrak{A}\leqq\mathfrak{B}$ if and only if $A\in\mathfrak{A}$ implies $A<B$ for some $B\in\mathfrak{B}$.

(3.2) THEOREM. *If X has an (invariant) $\mathfrak{M}$-cover then X has a maximal (invariant) $\mathfrak{M}$-cover.*

PROOF. Let $\{\mathfrak{A}_\lambda: \lambda\in\Lambda\}$ be a totally ordered subset of $\boldsymbol{A}_{\mathfrak{M}}$. For each $e\in X$ and $\lambda\in\Lambda$ let $A_{\lambda e}\in\mathfrak{A}_\lambda$ be such that $e\in A_{\lambda e}$. Put $A_e=\cup\{A_{\lambda e}: \lambda\in\Lambda\}$. $\mathfrak{N}_e=\{A_{\lambda e}: \lambda\in\Lambda\}$ is a tower in $\mathfrak{M}$, hence $A_e\in\mathfrak{M}$. If $E(A_e\cap A_{e'})\neq\square$, then $E(A_{\lambda e}\cap A_{\lambda' e'})\neq\square$ for some $\lambda, \lambda'\in\Lambda$. Without loss of generality $\mathfrak{A}_{\lambda'}\leqq\mathfrak{A}_\lambda$. Hence $A_{\lambda' e'}<A_{\lambda e'}$, and therefore by (3.1) (ii), $A_{\lambda e}=A_{\lambda e'}$. Hence $A_e=A_{e'}$. Thus $\mathfrak{A}=\{A_e: e\in X\}\in\boldsymbol{A}_{\mathfrak{M}}$, and $\mathfrak{A}\geqq\mathfrak{A}_\lambda$ for each $\lambda\in\Lambda$, whence it follows by Zorn's Lemma that $\boldsymbol{A}_{\mathfrak{M}}$ has a maximal element. The proof for maximal invariant covers is analogous.

For any graph X the *trivial* cover

$$\mathfrak{T}=\{(e): e\in E(X)\}\cup\{(x): x\in X, d(X; x)=0\}$$

is a complete invariant cover which refines every complete cover of X. Consequently, every graph has a maximal complete cover and a maximal invariant complete cover.

(3.3) DEFINITION. A graph X is called a *quasi-derivative* if and only if it is connected and has an invariant maximal complete cover (i.e., a maximal complete cover which is invariant).

If $\mathfrak{A}$ is any cover of X we define the *$\mathfrak{A}$-dual*[6] $\delta_{\mathfrak{A}}X$ of X as follows:

$$V(\delta_{\mathfrak{A}}X)=\mathfrak{A},$$

$$E(\delta_{\mathfrak{A}}X)=\{[A, B]: A, B\in\mathfrak{A},\ A\neq B,\ A\cap B\neq\square\}.$$

[5] A *tower* is a collection of subgraphs totally ordered by inclusion.

[6] $\delta_{\mathfrak{A}}X$ actually is the nerve of $\mathfrak{A}$; we have preferred the name dual in view of theorem (3.6).

(3.4) REMARK. *If X has no isolated vertices, then $\delta_{\mathfrak{T}} X = \partial X$.*

Put $G_{\mathfrak{A}} = \{\varphi \in G(X) : \varphi \mathfrak{A} = \mathfrak{A}\}$. Then $\varphi A_e = A_{\varphi e}$ for each $\varphi \in G_{\mathfrak{A}}$ and $e \in X$. It is clear from the definition of $\delta_{\mathfrak{A}} X$ that $\delta^*_{\mathfrak{A}} : G_{\mathfrak{A}} \to G(\delta_{\mathfrak{A}} X)$ given by

$$(\delta^*_{\mathfrak{A}} \varphi) A_e = A_{\varphi e}, \qquad \varphi \in G_{\mathfrak{A}}, \qquad A_e \in \delta_{\mathfrak{A}} X,$$

is a homomorphism. In general $\delta^*_{\mathfrak{A}}$ is neither one-one nor onto.

Let $\mathfrak{A}$ be a complete cover of X. A vertex $x \in X$ is called *regular* if and only if there exist two distinct members $A, B \in \mathfrak{A}$ such that $E(A) \neq \square \neq E(B)$, and $x \in A \cap B$.

(3.5) LEMMA. *Let $\mathfrak{A}$ be a complete cover of X. Then a n.a.s.c. that $\delta^*_{\mathfrak{A}}$ be a monomorphism is that each $A \in \mathfrak{A}$ contain at most one non-regular vertex.*

PROOF. Necessity. Let x, y be two distinct non-regular vertices of A. Then $V(X; x) \cup V(X; y) \subset V(A)$, and hence the transposition (x, y) belongs to $\operatorname{Ker} \delta^*_{\mathfrak{A}}$.

Sufficiency. Let $\varphi \in \operatorname{Ker} \delta^*_{\mathfrak{A}}$. Then $\varphi A = A$ for each $A \in \mathfrak{A}$. Let $x \in X$ be regular. Then there exist $A, B \in \mathfrak{A}$, $A \neq B$, such that $x = A \cap B$. Hence $\varphi x = \varphi A \cap \varphi B = A \cap B = x$. By hypothesis it then follows that for each $A \in \mathfrak{A}$, $\varphi | A$ is a permutation of the vertices of A which leaves all but at most one vertex fixed. Hence $\varphi | A = 1$, and therefore $\varphi = 1$.

If $\mathfrak{A}$ is a complete cover of X we denote by $\underline{X}$ the maximal subgraph of X, whose vertices are the regular vertices of X. For $x \in \underline{X}$ let D_x be the maximal subgraph of $\delta_{\mathfrak{A}} X$ with $V(D_x) = \{A \in \mathfrak{A} : x \in A,\ E(A) \neq \square\}$.

(3.6) THEOREM. *Let $\mathfrak{A}$ be a complete cover of X. Then $\mathfrak{D} = \{D_x : x \in \underline{X}\}$ is a complete cover of $\delta_{\mathfrak{A}} X$, and $\delta_{\mathfrak{D}} \delta_{\mathfrak{A}} X \approx \underline{X}$.*

PROOF. It is clear that D_x is a complete subgraph of $\delta_{\mathfrak{A}} X$. Let $\varepsilon = [A, B] \in E(\delta_{\mathfrak{A}} X)$, $A \cap B = x$. Then $A, B \in D_x$, hence $|D_x| \geqq 2$, and $\varepsilon \in D_x$. Hence $\cup \{D_x : x \in \underline{X}\} = \delta_{\mathfrak{A}} X$. If $[A, B] \in D_x \cap D_y$ then $x = A \cap B = y$ and hence $D_x = D_y$. Thus $\mathfrak{D}$ is a complete cover of $\delta_{\mathfrak{A}} X$.

To show that $\delta_{\mathfrak{D}} \delta_{\mathfrak{A}} X \approx \underline{X}$ define $\tau : \delta_{\mathfrak{D}} \delta_{\mathfrak{A}} X \to \underline{X}$ by

$$\tau D_x = x, \qquad D_x \in V(\delta_{\mathfrak{D}} \delta_{\mathfrak{A}} X) = \mathfrak{D}.$$

Clearly τ is well defined, one-one and onto. If $[D_x, D_y] \in E(\delta_{\mathfrak{D}} \delta_{\mathfrak{A}} X)$, then $D_x \cap D_y = A \in \mathfrak{A}$. Hence $x, y \in A$, and since A is complete, $[x, y] \in A$, whence $[x, y] \in \underline{X}$. Thus τ is a monomorphism. Reversal of the argument yields that τ^{-1} is likewise a monomorphism.

$\mathfrak{D}$ will be called the cover of $\delta_{\mathfrak{A}} X$ *induced* by $\mathfrak{A}$.

(3.7) DEFINITION. A complete cover $\mathfrak{A}$ of a graph X is called *uniform* if given any 3-circuit C of X there is an $A \in \mathfrak{A}$ such that $C \subset A$.

(3.8) THEOREM. *If X has a uniform cover $\mathfrak{A}$, then $\mathfrak{A}$ is the only maximal complete cover of X, and hence X is a quasi-derivative. Moreover, the induced cover $\mathfrak{D}$ of $\delta_{\mathfrak{A}} X$ is likewise uniform so that $\delta_{\mathfrak{A}} X$ is also a quasi-derivative.*

PROOF. Let $\mathfrak{B}$ be a complete cover of X and suppose there is a $B \in \mathfrak{B}$ such that $B \not\subset A$ for any $A \in \mathfrak{A}$. Then $|B| \geqq 3$ and there exist at least two

distinct members A_0, A_1 of $\mathfrak{A}$ such that $E(A_i \cap B) \neq \square$, $i=0,1$. Let $e_i = [x_i, y_i] \in A_i \cap B$, $i=0,1$. Then $e_0 \neq e_1$. Hence w.l.o.g. $x_0 \neq x_1$. Also, we may assume that if e_0 and e_1 are adjacent, then $y_0 = y_1$. It follows that $C_i = (x_i, y_i, x_{1-i})$ is a 3-circuit of B. Note that if $C_i < A_{(i)} \in \mathfrak{A}$ then $A_{(i)} = A_i$, $i=0,1$. $[x_0, x_1] \in C_0 \cap C_1$, hence if C_0 and C_1 each belong to some member of $\mathfrak{A}$ then $A_0 = A_1$, a contradiction. It follows that either C_0 or C_1 violates the hypothesis of uniformity.

Suppose that (A, B, C) is a 3-circuit of $\delta_{\mathfrak{A}} X$ which is not contained in any member of $\mathfrak{D}$. Put $A \cap B = x$, $B \cap C = y$, $C \cap A = z$. Then x, y, z are three distinct vertices of X, and since $[x, y] \in B$, $[y, z] \in C$, $[z, x] \in A$, (x, y, z) does not belong to any member of $\mathfrak{A}$ contrary to the uniformity of $\mathfrak{A}$.

4. Even graphs

The importance of complete covers (and quasi-derivatives) lies in their close relationship to a certain class of even graphs. In describing this connection it will be convenient to use the following notation. For a given graph X we define X^2 by

$$V(X^2) = V(X),$$
$$E(X^2) = E(X) \circ E(X) - \Delta,$$

where $\Delta = \{[x, x] : x \in X\}$. That is, $[x, y] \in X^2$ if and only if there is a $z \in X$ such that $[x, z]$, $[z, y] \in X$, and $x \neq y$.

Now let X be a connected even graph. Then $V(X)$ consists of exactly two colour classes M_0, M_1 such that $x, y \in M_\alpha$ implies $[x, y] \notin X$, $\alpha = 0, 1$. Hence X^2 has exactly two components X_0, X_1 given by $V(X_\alpha) = M_\alpha$, $\alpha = 0, 1$. We shall show that with suitable restrictions on X, X_0 and X_1 are quasi-derivatives.

We shall use the following terminology and notation: An even graph X will be called *admissible* if it is connected, and mesh $X \geqq 6$. For a connected even graph X we denote by

(1) Y_α the maximal subgraph of X_α whose vertices are of degree $\geqq 2$ in X, $\alpha = 0, 1$;

(2) $X^{(\alpha)}$ the maximal subgraph of X with $V(X^{(\alpha)}) = V(Y_\alpha) \cup M_{1-\alpha}$, $\alpha = 0, 1$;

(3) N_α the set of all vertices of M_α which belong to some 6-circuit of X, $\alpha = 0, 1$;

(4) $A(x)$ the maximal subgraph of X^2 with $V(A(x)) = V(X; x)$, $x \in X$.

(4.1) LEMMA. *Let X be an admissible even graph. If $|X_\alpha| \geqq 2$, then $\mathfrak{A}_\alpha = \{A(y) : y \in Y_{1-\alpha}\}$ is a complete cover of X_α, and $\delta_{\mathfrak{A}_\alpha} X_\alpha \approx Y_{1-\alpha}$, $\alpha = 0, 1$.*

PROOF. That each $A(y)$ is complete is clear. Since $d(X; y) \geqq 2$, $E(A(y)) \neq \square$. Let $[x, z] \in X_\alpha$. Then there is a $y \in M_{1-\alpha}$ such that $[x, y]$, $[y, z] \in X$. Hence $y \in Y_{1-\alpha}$, and $[x, z] \in A(y)$. Therefore $X_\alpha = \cup\, \mathfrak{A}_\alpha$. If $E(A(y) \cap A(y')) \neq \square$, $[x, z] \in A(y) \cap A(y')$, say, then $y = y'$. Otherwise (x, y, z, y') is a 4-circuit of X, and hence mesh $X = 4$, a contradiction.

Define $\tau_\alpha: \delta_{\mathfrak{A}_\alpha} X_\alpha \to Y_{1-\alpha}$ by

$$\tau_\alpha A(y) = y, \qquad A(y) \in V(\delta_{\mathfrak{A}_\alpha} X_\alpha) = \mathfrak{A}_\alpha.$$

Clearly τ_α is well defined, one-one and onto. If $[A(y), A(y')] \in \delta_{\mathfrak{A}_\alpha} X_\alpha$ there is a $z \in A(y) \cap A(y')$, hence $[y, z], [z, y'] \in X$, and since $y \neq y'$, $[y, y'] \in X^2$, i.e., $[y, y'] \in Y_{1-\alpha}$. Thus τ_α is a monomorphism, and, since the preceding argument can be reversed, an isomorphism.

REMARK. If $|X_\alpha| = 1$ we will agree to denote by $\mathfrak{A}_\alpha$ the cover consisting of X_α alone.

Lemma (4.1) has the following converse.

(4.2) LEMMA. *Let X be a connected graph, $\mathfrak{A}$ a complete cover of X. Define $X_{\mathfrak{A}}$ by*

$$V(X_{\mathfrak{A}}) = V(X) \cup V(\delta_{\mathfrak{A}} X),$$

$$E(X_{\mathfrak{A}}) = \{[x, A] : x \in V(X), A \in \mathfrak{A}, x \in A\}.$$

Then $X_{\mathfrak{A}}$ is an admissible even graph, and $X^2_{\mathfrak{A}} = X + \delta_{\mathfrak{A}} X$.

PROOF. $V(X)$ and $\mathfrak{A}$ are disjoint. Hence $X_{\mathfrak{A}}$ is even, $V(X)$ and $\mathfrak{A}$ being the two colour classes of $X_{\mathfrak{A}}$.

Let $x, y \in X$, $x \neq y$, and let $(x_0, \ldots, x_n)$ be a path of X joining x and y, i.e., $x_0 = x$, $x_n = y$. Put $[x_{i-1}, x_i] = e_i$, $i = 1, \ldots, n$. Then $x_{i-1}, x_i \in A_{e_i} \in \mathfrak{A}$, i.e., $[x_{i-1}, A_{e_i}], [A_{e_i}, x_i] \in X_{\mathfrak{A}}$. Hence $(x_0, A_{e_1}, x_1, \ldots, x_{n-1}, A_{e_n}, x_n)$ is a path of $X_{\mathfrak{A}}$ joining x and y. Since any vertex of $X_{\mathfrak{A}}$ either belongs to $V(X)$ or is adjacent to a vertex of X it follows that $X_{\mathfrak{A}}$ is connected.

If $X_{\mathfrak{A}}$ has a 4-circuit (x, A, y, B), then $x, y \in A \cap B$, and since A and B are complete, $[x, y] \in A \cap B$, whence $A = B$, a contradiction. Thus $X_{\mathfrak{A}}$ is admissible.

Let X_0, X_1 be the maximal subgraphs of $X^2_{\mathfrak{A}}$ with $V(X_0) = V(X)$, $V(X_1) = \mathfrak{A}$, respectively. $[x, y] \in X_0$ implies the existence of an $A \in \mathfrak{A}$ such that $x, y \in A$. By the completeness of A, $[x, y] \in A < X$. Conversely, if $e = [x, y] \in X$, then $x, y \in A_e \in \mathfrak{A}$, and hence $[x, y] \in X_0$. Thus $X_0 = X$. Similarly it follows that $X_1 = \delta_{\mathfrak{A}} X$.

(4.3) LEMMA. *Let X be an admissible even graph such that* (i) $d(X; x) < \infty$ *for each $x \in N_\alpha$,* (ii) $d(X; y) \geqq 3$ *for each $y \in N_{1-\alpha}$. Then $\mathfrak{A}_\alpha$ is a maximal complete cover of X_α.*

PROOF. Let $\mathfrak{B}_\alpha$ be a complete cover of X_α with $\mathfrak{B}_\alpha > \mathfrak{A}_\alpha$. Then there is a $y \in Y_{1-\alpha}$ and a $B \in \mathfrak{B}_\alpha$ such that $A(y)$ is properly contained in B. Let $z \in V(B) - V(A(y))$. Since B is complete, $[x, z] \in B$ for each $x \in A(y)$. Hence for each $x \in A(y)$ there is a vertex $w_x \in V(X; x) \cap V(X; z) < M_{1-\alpha}$, and $w_x = w_{x'}$ if and only if $x = x'$; otherwise mesh $X = 4$. Since $y \in Y_{1-\alpha}$, $d(X; y) \geqq 2$. Hence for each $x \in A(y)$ there exists an $x' \in A(y)$ such that $(y, x, w_x, z, w_{x'}, x')$ is a 6-circuit of X. By condition (ii), $d(X; w_x) \geqq 3$, $x \in A(y)$. Let $z_x \in V(X; w_x) - \{x, z\}$. Since mesh $X > 4$ no two z_x's are equal; for the same reason $z_x \neq w_{x'}$ for every $x, x' \in A(y)$. $[x, z] \in A(w_x) \cap B$, and since $\mathfrak{B}_\alpha > \mathfrak{A}_\alpha$, $[x, z] \in B_x \cap B$, where $A(w_x) < B_x \in \mathfrak{B}_\alpha$. Hence $B_x = B$, so that $z_x \in A(w_x) < B$. By the com-

pleteness of B, $[z_x, x'] \in B$ for all $x, x' \in A(y)$. Hence there exist vertices $w_{xx'} \in V(X; x) \cap V(X; z_{x'})$, $x, x' \in A(y)$, and since mesh $X \geqq 6$, $w_{xx} = w_x$. We also note that if $x \neq x'$, $w_{xx'}$ is distinct from all vertices previously defined, and that no two $w_{xx'}$'s are equal. Each $w_{xx'}$ belongs to a 6-circuit of X, hence by (ii), $d(X; w_{xx'}) \geqq 3$. We let $z_{xx'} \in V(X; w_{xx'}) - \{x, z_{x'}\}$, and repeat the preceding argument obtaining vertices $w_{xx'x''}$, $z_{xx'x''}$, etc.

It follows that $w_x, w_{xx'}, w_{xx'x''}, \ldots \in V(X; x)$; hence $d(X; x)$ is infinite, contrary to hypothesis (i).

(4.4) Remark. Lemma (4.3) becomes false if condition (ii) is omitted. One can, however, show the following: *If X is an admissible even graph such that $d(X; x) \leqq 2$ for each $x \in N_\alpha$, then either $\mathfrak{A}_\alpha$ is a maximal complete cover of X_α, or X is a 6-circuit, and in that case $\mathfrak{A}_\alpha$ is not maximal. On the other hand, if $n \geqq 3$, there exist infinitely many non-isomorphic admissible even graphs X such that* (i) *$d(X; x) \leqq n$ for each $x \in N_\alpha$, with equality for at least one x, and* (ii) *$\mathfrak{A}_\alpha$ is not maximal.*

Proof. Using the notation of the proof of (4.3) we have that $2 \leqq d(X; y) \leqq d(X; z) \leqq 2$. Also, by hypothesis, $d(X; x) = 2$, $x \in A(y)$. Hence either X is a 6-circuit or $d(X; w_x) \geqq 3$ for some $x \in A(y)$. But then $w_x \in V(X; x')$, $x' \in A(y)$, hence $d(X; x') \geqq 3$, a contradiction.

For $n = 3$ an admissible even graph X for which $\mathfrak{A}_\alpha$ fails to be maximal, and which is not a 6-circuit, is given in Fig. 1. M_α is that colour class which contains the vertex x.

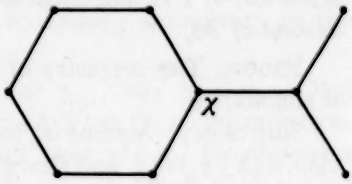

Fig. 1

By concatenation of any finite number of copies of X infinitely many such graphs can be obtained.

We shall now derive conditions for the invariance of $\mathfrak{A}_\alpha$.

Let $G^{(\alpha)}$ be the subgroup of $G(X^{(\alpha)})$ which leaves $M_{1-\alpha}$ invariant. For $\varphi \in G^{(\alpha)}$ define $\varphi_{1-\alpha} = \varphi | M_{1-\alpha}$. Let $[x, y] \in X_{1-\alpha}$. Then there exists a $z \in Y_\alpha$ such that $[x, z], [z, y] \in X^{(\alpha)}$. Hence $[\varphi x, \varphi z], [\varphi z, \varphi y] \in X^{(\alpha)}$, and therefore $[\varphi_{1-\alpha} x, \varphi_{1-\alpha} y] \in X_{1-\alpha}$. Thus $\varphi_{1-\alpha} \in G(X_{1-\alpha})$. It follows that $f_\alpha : G^{(1-\alpha)} \to G(X_\alpha)$ defined by $f_\alpha \varphi = \varphi_\alpha$, $\varphi \in G^{(1-\alpha)}$, is a homomorphism.

(4.5) Lemma. *Let X be an admissible even graph. Then $f_\alpha : G^{(1-\alpha)} \to G(X_\alpha)$ is a monomorphism, $\alpha = 0, 1$.*

Proof. Let $\varphi \in \operatorname{Ker} f_\alpha$. Then $\varphi x = x$ for each $x \in M_\alpha$. Let $y \in Y_{1-\alpha}$. Then $V(X^{(1-\alpha)}; y) \subset M_\alpha$. Hence $V(X; \varphi y) = V(X^{(1-\alpha)}; \varphi y) = \varphi V(X^{(1-\alpha)}; y) = V(X^{(1-\alpha)}; y) = V(X; y)$. Since $d(X; y) = d(X; \varphi y) \geqq 2$, and mesh $X > 4$, $\varphi y = y$. Hence $\varphi = 1$.

(4.6) Lemma. *Let X be an admissible even graph, $|X_\alpha| \geqq 2$. Then a n.a.s.c. that $\mathfrak{A}_\alpha$ be invariant is that $f_\alpha : G^{(1-\alpha)} \approx G(X_\alpha)$*

Proof. Suppose $\mathfrak{A}_\alpha$ is invariant. In view of lemma (4.5) it suffices to show that f_α is onto. Let $\tau \in G(X_\alpha)$, $x \in Y_{1-\alpha}$. By the invariance of $\mathfrak{A}_\alpha$ there is a $w_x \in Y_{1-\alpha}$ such that $A(w_x) = \tau A(x)$. w_x is unique, otherwise mesh $X = 4$.

Define $\varphi: X^{(1-\alpha)} \to X^{(1-\alpha)}$ by

$$\varphi x = \begin{cases} \tau x, & x \in M_\alpha \\ w_x, & x \in Y_{1-\alpha}. \end{cases}$$

Let $[x, y] \in X^{(1-\alpha)}$, $x \in Y_{1-\alpha}$, $y \in M_\alpha$. Then $[\varphi x, \varphi y] = [w_x, \tau y]$. By definition of w_x, $y \in A(x)$ implies $\tau y \in A(w_x)$; hence $[w_x, \tau y] \in X^{(\alpha-1)}$, i.e., $\varphi \in G^{(1-\alpha)}$.

Suppose f_α is an isomorphism. For given $\sigma \in G(X_\alpha)$ put $\tau = (f_\alpha^{-1} \sigma) | Y_{1-\alpha}$. Then $\sigma \mathfrak{A}_\alpha = \{\sigma A(y) : y \in Y_{1-\alpha}\} = \{A(\tau y) : y \in Y_{1-\alpha}\} = \mathfrak{A}_\alpha$. Hence $\mathfrak{A}_\alpha$ is invariant.

Let $e \in X_\alpha$. By $\boldsymbol{K}_e$ denote the set of all complete subgraphs of X_α which contain e. Put $\sup\{|K| : K \in \boldsymbol{K}_e\} = m(e)$, and $\boldsymbol{M}_e = \{K \in \boldsymbol{K}_e : |K| = m(e)\}$. e will be called *exceptional* if and only if $m(e) = 3$ and $|\boldsymbol{M}_e| = 1$; it will be called *singular* if and only if $m(e) = 3$ and $|\boldsymbol{M}_e| \geqq 2$.

By $\overline{C}_4$ denote the graph obtained by dividing each edge of the complete 4-graph C_4 by a new vertex.

Let $e \in X_\alpha$. By y_e we shall denote the unique vertex of $Y_{1-\alpha}$ such that $e \in A(y_e)$.

(4.7) LEMMA. *Let X be an admissible even graph with the property that if $X > Y \approx \overline{C}_4$, then $d(Y; x) = 2$ for each $x \in M_\alpha \cap V(Y)$, and let $e \in X_\alpha$. Then a n.a.s.c. that $\boldsymbol{M}_e$ have a unique maximal element is that e be non-singular. Moreover, if e is non-singular and non-exceptional, then $A(y_e)$ is the maximal element of $\boldsymbol{M}_e$.*

PROOF. The necessity of the condition is immediate from the definition of singularity.

Sufficiency: Assume e_0 to be non-singular. If $m(e_0) = 2$ then $\boldsymbol{K}_{e_0} = \{(e_0)\}$. Also, $d(X; y_{e_0}) = 2$, and hence $A(y_{e_0}) = (e_0)$. If $m(e_0) = 3$, then e_0 is exceptional, and the assertion follows at once from the definition of exceptionality. If $m(e_0) \geqq 4$ let $K \in \boldsymbol{K}_{e_0}$ with $|K| = 4$ be given. Since mesh $X > 4$ only two possibilities can arise. Either $y_e = y_{e_0}$ for each $e \in K$, and then $K < A(y_{e_0})$, or else all y_e's are distinct. But in the latter case the graph Y with

$$V(Y) = V(K) \cup \{y_e : e \in K\},$$
$$E(Y) = \{[x, y_e] : x \in K, e \in K, e \in E(K; x)\},$$

is isomorphic to $\overline{C}_4$ and violates the hypothesis of the lemma.

A vertex $x \in X_\alpha$ will be called *singular* if it belongs to a 3-circuit C such that either C contains a singular edge or all three edges of C are exceptional. If x belongs to a 3-circuit with edges e_1, e_2, e_3 such that e_1, e_2 are exceptional, e_3 is singular, and $x = e_1 \cap e_2$, then x is called *exceptional*. Note that every exceptional vertex is singular.

By Q we denote the graph with $V(Q) = \{0, 1, 2, 3\}$, $E(Q) = \{[0, i] : i = 1, 2, 3\} \cup \{[1, 2]\}$.

(4.8) THEOREM. *Let X be an admissible even graph such that* (i) *X satisfies the hypothesis of lemma* (4.7), *and* (ii) *$d(X; x) \leqq 2$ for each singular vertex $x \in X_\alpha$. Suppose that X_α is not isomorphic to C_3, ∂Q, $\partial^2 Q$, ∂C_4, and has a complete cover $\mathfrak{A}$ such that* (1) *if e is a non-exceptional edge of X_α then A_e is*

a maximal element of $\boldsymbol{M}_e$; (2) *if* $e \in X_\alpha$ *is exceptional,* $\boldsymbol{M}_e = \{C\}$, *then* $A_e = C$ *if all edges of* C *are exceptional*; (3) *if* x *is a singular or exceptional vertex of* X_α *then* x *belongs to at most two distinct members of* $\mathfrak{A}$. *Then* $\mathfrak{A} = \mathfrak{A}_\alpha$.

PROOF. Let $e \in X_\alpha$. We must show that $A_e = A(y_e)$. If e is non-singular and non-exceptional this follows from lemma (4.7). It thus remains to consider two cases.

Case (i): $e = [x, y]$ is singular. Let $C_i = (x, y, x_i)$, $i = 0, 1$, be two distinct members of $\boldsymbol{M}_e$. Put $e_{xi} = [x, x_i]$, $e_{yi} = [y, x_i]$, $i = 0, 1$. Then, without loss of generality, $y_e = y_{e_{x0}} = y_{e_{y0}}$, while y_e, $y_{e_{x1}}$, $y_{e_{y1}}$ are distinct. This follows from the fact that the members of each triple $(y_e, y_{e_{xi}}, y_{e_{yi}})$, $i = 0, 1$, are either all equal or all distinct. If $y_e = y_{e_{xi}} = y_{e_{yi}}$, $i = 0, 1$, then $m(e) \geqq 4$, a contradiction against the singularity of e; if all five y's are distinct, then $d(X; x) \geqq 3$, contrary to hypothesis (ii).

Now assume that $A_e \neq A(y_e)$, i.e., $A_e = C_1$. Note that $P = C_0 \cup C_1 \approx \partial Q$. Hence $E(X_\alpha) - E(P) \neq \square$. This implies that $E = E(X_\alpha; x) \cup E(X_\alpha; y) - E(P) \neq \square$. For if $E = \square$, then e_{x0} and e_{y0} are exceptional, and $A_{e_{x0}} = e_{x0}$, $A_{e_{y0}} = e_{y0}$. Consequently, x_0 is an exceptional vertex of X_α, and by (3), $A_{e_{x0}}$ and $A_{e_{y0}}$ are the only members of A containing x_0, i.e., $V(X_\alpha; x_0) = \{x, y\}$. It follows that there is an $e' \in E(X_\alpha; x_1) - E(P)$. If $e' = [x_1, w]$, then $w \notin P$. Otherwise $w = x_0$, and then $P \cup e'$ is a complete 4-graph containing e, contrary to the singularity of e.

Since $d(X; x) = 2$ it follows that $y_{e'} = y_{e_{x1}}$ or $y_{e_{y1}}$, say $y_{e'} = y_{e_{x1}}$. But then $[x, w] \in A(y_{e_{x1}})$, and since $w \notin P$, $E(X_\alpha; x) - E(P) \neq \square$, a contradiction. Hence there exists an edge $e' = [x, y'] \in X_\alpha$ such that $y' \notin P$. By the same argument as before, $y_{e'} = y_{e_{x1}}$. Note that $d(X; y_{e'}) = 3$, for if $d(X; y_{e'}) \geqq 4$ then $A_{e'} = A(y_{e'})$, and hence $e_{x1} \in A_{e'} \cap A_e$, a contradiction. From $y_{e'} = y_{e_{x1}}$ it also follows that $[x_1, y'] \in E(X_\alpha)$.

Now consider $[x_0, y']$. If $[x_0, y'] \notin X_\alpha$, then C_1, A_{e_x} and $A_{e'}$ are three distinct members of $\mathfrak{A}$ containing x, contrary to hypothesis (3). Hence $[x_0, y'] \in X_\alpha$. As none of the four edges e, e', e_{x0}, e_{x1} is contained in a complete 4-graph, it follows that the maximal subgraph R of X_α with vertices x, y, x_0, x_1, y' is isomorphic to $\partial^2 Q$. By hypothesis there is a $z \in V(X_\alpha) - V(R)$ adjacent to some vertex of R.

If $[x, z] \in X_\alpha$ then either $d(X; y_e) \geqq 4$ or $d(X; y_{e'}) \geqq 4$ or $d(X; x) \geqq 3$, none of which can happen. Hence z is not adjacent to x; without loss of generality we may assume that it is adjacent to y. Put $[y, z] = e''$. Since y is singular and $d(X; y) = 2$, $y_{e''} = y_{e_{y1}}$ and $d(X; y_{e''}) = 3$. By an argument of the same type as in the preceding paragraphs one then shows that the maximal subgraph S of X_α with $V(S) = \{x, y, z, x_0, x_1, y'\}$ is isomorphic to ∂C_4, and that that part of $\mathfrak{A}$ which covers S consists of C_1, (x, x_0, y'), (x_0, y, z) and (x_1, y', z). Since $X_\alpha \not\approx \partial C_4$, there exists a $w \in V(X_\alpha) - V(S)$ adjacent to some vertex of S. $G(\partial C_4)$ acting transitively on $V(\partial C_4)$ we may assume that $e_w = [w, x] \in X_\alpha$. $d(X; x) \leqq 2$ implies $y_{e_w} = y_e$ or $y_{e'}$, say $y_{e_w} = y_e$, but then $d(X; y_e) \geqq 4$, a contradiction. Hence $A_e = A(y_e)$ as was to be shown.

Case (ii): $e=[x, y]$ is exceptional. Let $C=(x, y, z)$ be the unique 3-circuit containing e, and put $[x, z]=e_x$, $[y, z]=e_y$. $X_\alpha \not\approx C_3$ implies that there is a vertex $w \in V(X_\alpha)-V(C)$ which is adjacent to some vertex of C.

If e_x (or e_y) is non-exceptional then by the earlier parts of the proof $A_{e_x}=A(y_{e_x})$. If $y_{e_x}=y_e$, then $e \in A(y_e)=A(y_{e_x})=A_{e_x}$, and hence $A_e=A(y_e)$. If $y_{e_x} \neq y_e$, then $d(X; y_e)=2$, hence $A(y_e)=e$. Also, since $A_{e_x}=A(y_{e_x}) \neq C$, $A_e=e$. Hence again $A_e=A(y_e)$.

We may therefore suppose that the three edges e, e_x, e_y are exceptional. Without loss of generality $e'=[w, x] \in X_\alpha$. There are now two possibilities. Either y_e, y_{e_x}, y_{e_y} are distinct or $y_e=y_{e_x}=y_{e_y}$. In the first case the fact that x is singular implies that $y_{e'}=y_e$ or y_{e_x}, say $y_{e'}=y_e$. But then $w, x, y \in V(X; y_e)$ and hence (w, x, y) is a 3-circuit containing e so that e is non-exceptional, a contradiction. It follows that $y_e=y_{e_x}=y_{e_y}$, i.e., $A(y_e)=C$. But by hypothesis (2), $A_e=C$ whence $A_e=A(y_e)$ in this case also.

(4.9) COROLLARY. *If X, X_α and $\mathfrak{A}_\alpha$ satisfy the hypotheses of theorem* (4.8), *then $\mathfrak{A}_\alpha$ is maximal and invariant, and hence X_α is a quasi-derivative with respect to $\mathfrak{A}_\alpha$.*

5. Derivatives

We shall now derive Whitney's theorem from theorem (4.8). Let X be any connected graph with $E(X) \neq \square$. Let $\mathfrak{T}$ be the trivial cover of X and put $\overline{X}=X_{\mathfrak{T}}$ (notation of lemma (4.2)). That is to say,

$$V(\overline{X}) = V(X) \cup E(X),$$
$$E(\overline{X}) = \{[e, x] : e \in E(X; x)\}.$$

By (4.2), $\overline{X}$ is an admissible even graph, and if we denote by $\overline{X}_0$ and $\overline{X}_1$ the two components of $\overline{X}^2$ with $V(\overline{X}_0)=V(X)$ and $V(\overline{X}_1)=E(X)$, respectively, then $\overline{X}_0=X$ and $\overline{X}_1=\delta_{\mathfrak{T}} X=\partial X$. Note that if $e=[x, y] \in E(X)$ then $V(\overline{X}; e)=\{x, y\}$, hence $d(\overline{X}; e)=2$ for each $e \in E(X)$.

Let $\mathfrak{E}$ be the complete cover of ∂X induced by the trivial cover of X. $\mathfrak{E}$ will henceforth be called the *natural* cover of ∂X. If we denote by $\underline{X}$ the maximal subgraph of X with $V(\underline{X})=\{x \in X : d(X; x) \geqq 2\}$, (i.e., the subgraph of X whose vertices are regular w.r.t. the trivial cover), then $\mathfrak{E}=\{E_x : x \in \underline{X}\}$, where E_x denotes the maximal subgraph of ∂X with $V(E_x)=E(X; x)$. For $\varepsilon \in E(\partial X)$ we shall denote by E_ε that member of $\mathfrak{E}$ which contains ε. $\mathfrak{E}$ has the following properties:

(5.1) (i) *If $e=[x, y] \in V(\partial X)$ and if $d(X; x) \geqq 2$, $d(X; y) \geqq 2$, then E_x and E_y are the only members of $\mathfrak{E}$ containing e.*

(ii) *If $e=[x, y] \in V(\partial X)$ and if $d(X; x) \geqq 2$, $d(X; y) \neq 1$, then E_x is the only member of $\mathfrak{E}$ containing e. Thus e is regular with respect to $\mathfrak{E}$ if and only if both ends of e are of degree $\geqq 2$.*

(iii) *If $\varepsilon_1, \varepsilon_2, \varepsilon_3$ are three exceptional edges of ∂X forming a 3-circuit C of ∂X, then either $\partial X=C$ or $E_{\varepsilon_i}=C$, $i=1, 2, 3$.*

PROOF of (iii). If $\partial X \neq C$ there is an $\varepsilon_0 \in \partial X$ adjacent to both ε_1 and ε_2, say. If $E_{\varepsilon_i} \neq C$, then $E_{\varepsilon_i} = \varepsilon_i$, $i = 1, 2$, and then the vertex common to $\varepsilon_0, \varepsilon_1, \varepsilon_2$ belongs to three distinct members of $\mathfrak{E}$, viz. $E_{\varepsilon_0}, (\varepsilon_1), (\varepsilon_2)$, a contradiction.

It follows that if ∂X is not isomorphic to C_3, ∂Q, $\partial^2 Q$, or ∂C_4, then theorem (4.8) is applicable to $\overline{X}$ for $\alpha = 1$ with the result that $\mathfrak{A}_1 = \mathfrak{E}$ is maximal, and the only complete cover of ∂X with the property that each vertex of ∂X belongs to at most two members of $\mathfrak{E}$. As a consequence, if X and X' are two connected graphs subject to the above restrictions, if $\mathfrak{E}$ and $\mathfrak{E}'$ are the natural covers of ∂X and $\partial X'$, respectively, and if $\eta : \partial X \approx \partial X'$, then $\mathfrak{E}' = \eta\,\mathfrak{E}$, and hence $\delta_{\mathfrak{E}} \partial X \approx \delta_{\mathfrak{E}'} \partial X'$, i.e., $\underline{X} \approx \underline{X}'$. This is easily used to prove the following theorem.

(5.2) THEOREM. *Let X and X' be two non-empty connected graphs neither of which is isomorphic to C_3. Then $\partial X \approx \partial X'$ implies $X \approx X'$.*

PROOF. It is easily shown that the theorem holds for the exceptional graphs C_4, Q, ∂Q (C_3 is excluded by hypothesis). In fact, as regards C_4, one can show that it is the only connected graph whose derivative has only singular edges. We omit the proof.

Suppose then that X is none of the exceptional graphs C_4, Q, ∂Q. If $d(X; x) = 1$ denote by e_x the edge of X incident with x. Then e_x is a non-regular vertex of ∂X; hence ηe_x is a non-regular vertex of $\partial X'$. Thus there exists a unique vertex $x'_{\eta e_x} \in X'$ such that $E(X'; x'_{\eta e_x}) = \{\eta e_x\}$.

Let $\sigma : \underline{X} \approx \underline{X}'$ be the isomorphism induced by η. σ can be extended to an isomorphism $\tau : X \approx X'$ by setting

$$\tau x = \begin{cases} \sigma x, & \text{if } d(X; x) \geqq 2 \\ x'_{\eta e_x}, & \text{if } d(X; x) = 1. \end{cases}$$

This completes the proof.

(5.3) THEOREM. *Let X be a connected graph not isomorphic to C_2, C_4, Q, ∂Q. Then $\partial^* : G(X) \to G(\partial X)$ given by*

$$\partial^* \varphi = \tilde{\varphi}, \qquad \varphi \in G(X),$$

is an isomorphism.

PROOF. If $|X| \leqq 1$, then $\partial X = \square$, and hence $G(X) = G(\partial X) = 1$. If $|X| = 2$, i.e., if $X \approx C_2$, then $|\partial X| = 1$, hence $G(\partial X) = 1$. If $X \approx C_4$, Q, ∂Q, then ∂^* is a monomorphism, and Im ∂^* is of index 2 in $G(\partial X)$.

If X is none of these graphs then by (4.5), $f_1 : \overline{G}^{(0)} \approx G(\partial X)$, where $\overline{G}^{(0)}$ is the subgroup of $G(\overline{X}^{(0)})$ which leaves $\overline{M}_1 = E(X)$ invariant, and $\overline{X}^{(0)}$ is the maximal subgraph of $\overline{X}$ with $V(\overline{X}^{(0)})$ consisting of $E(X)$ together with all vertices of X which are of degree $\geqq 2$.

$\overline{X}$ is obtained from X by subdividing each edge by a new vertex. Hence either (i) $h : G(X) \approx G(\overline{X})$, where $(h\varphi)x = \varphi x$, $x \in V(X)$, $(h\varphi)e = \tilde{\varphi} e$, $e \in E(X)$, or (ii) X is an n-circuit.

In case (i), $g : G(\overline{X}) \to \overline{G}^{(0)}$ given by $g\psi = \psi | \overline{X}^{(0)}$, $\psi \in G(\overline{X})$, is an isomorphism. Every $\psi \in G(\overline{X})$ leaves $E(X)$ invariant; hence so does $g\psi$. Let $\lfloor e, x \rceil \in \overline{X}^{(0)}$.

Then $d(X;x)\geqq 2$, $e\in E(X;x)$. Hence $\tilde{\varphi}e\in E(X;\varphi x)$, $d(X;\varphi x)\geqq 2$, where $\varphi=h^{-1}\psi$, and therefore $(g\psi)[e,x]=[\tilde{\varphi}e,\varphi x]\in\overline{X}^{(0)}$. Hence $g\psi\in\overline{G}^{(0)}$.

Let $\psi\in\operatorname{Ker} g$, $\varphi=h^{-1}\psi$. Then $\tilde{\varphi}e=e$ for each $e\in E(X)$, and $\varphi x=x$ for each $x\in X$ with $d(X;x)\geqq 2$. Now suppose $d(X;x)=1$. In the notation of the proof of (5.2), $E(\overline{X};x)=\{[e_x,x]\}$. Hence $E(\overline{X};\psi x)=\{[\tilde{\varphi}e_x,\varphi x]\}=\{[e_x,\psi x]\}$, and therefore $\psi x=x$. Thus g is a monomorphism. To see that g is onto let $\tau\in\overline{G}^{(0)}$ be given and define $\psi\in G(\overline{X})$ by

$$\psi e=\tau e,\quad \text{if}\quad e\in E(X)$$

$$\psi x=\begin{cases}\tau x, & \text{if}\quad d(X;x)\geqq 2,\quad x\in V(X)\\ x_{\tau e_x}, & \text{if}\quad d(X;x)=1,\quad x\in V(X).\end{cases}$$

Combining the three isomorphisms f_1, g, h we obtain that $f_1 g h\colon G(X)\approx G(\partial X)$, and $f_1 g h=\partial^*$.

In case (ii), $\overline{X}=\overline{X}^{(0)}$ is a $2n$-circuit, and, with h as defined in case (i), $h\colon G(X)\approx\overline{G}^{(0)}$. Hence $f_1 h\colon G(X)\approx G(\partial X)$, and $f_1 h=\partial^*$.

6. Representation of graphs by joins of complete graphs

In this section we slightly modify and thereby extend to infinite graphs a device which was introduced by IZBICKI [1] strictly for finite graphs. We will show that, in the terminology of [1], every graph has a minimal V-representation (lemma (6.5)).

(6.1) DEFINITION. Let Y be a graph, $\{X_y: y\in Y\}$ a collection of graphs indexed by $V(Y)$. By the *Y-join* of the X_y's is meant the graph $X=\vee(X_y;Y)$ given by

$$V(X)=\{(x,y): x\in X_y,\ y\in Y\},$$

$$E(X)=\{[(x,y),(x',y')]: [y,y']\in Y,\ \text{or}\ y=y'\ \text{and}\ [x,x']\in X_y\}.$$

Note that if $X_y\approx X_0$ for each $y\in Y$, then $X\approx Y\circ X_0$, the lexicographic product of Y and X_0 ([3], definition 2).

We shall be concerned with the case where Y is connected and each X_y is complete, i.e., $X_y\approx C_{n_y}$, $n_y\geqq 1$.

(6.2) DEFINITION. By a *representation* of a graph X is meant a collection of graphs $R=(Y;C_{n_y}: y\in Y)$, where C_{n_y} is the complete n_y-graph, $n_y\geqq 1$, and $X\approx\vee(C_{n_y};Y)$. Two representations $R=(Y;C_{n_y}: y\in Y)$ and $S=(Z;C_{m_z}: z\in Z)$ of X are *isomorphic* if and only if $Y\approx Z$. R and S are *equivalent* if and only if $\eta\colon Y\approx Z$, and $m_{\eta y}=n_y$ for each $y\in Y$. R is *trivial* if $n_y=1$ for each $y\in Y$.

Note that any two trivial representations of X are equivalent.

(6.3) DEFINITION. Let X be a graph. Define an equivalence relation on $V(X)$ by setting $x\sim y$ if and only if $V(X;x)\cup\{x\}=V(X;y)\cup\{y\}$. For each $x\in X$ put $x_*=\{y\in X: y\sim x\}$, and by X_* denote the graph with

$$V(X_*)=\{x_*: x\in X\},$$

$$E(X_*)=\{[x_*,y_*]: x_*\neq y_*,\quad [x,y]\in E(X)\}.$$

(6.4) LEMMA. *Let X be a graph. Then* (i) *there is a subgraph Y of X such that $Y \approx X_*$;* (ii) $X_{**} \approx X_*$; (iii) $X_* \approx X$ *if and only if* $x_* = \{x\}$ *for each* $x \in X$; (iv) $\sim$ *is invariant under isomorphisms.*

PROOF. $[x_*, y_*] \in X_*$ implies $[x, y] \in X$. For by the definition of X_*, $[x', y'] \in X$ for some $x' \in x_*$, $y' \in y_*$. Hence $y' \in V(X; x') \subset V(X; x') \cup \{x'\} = V(X; x) \cup \{x\}$, and since $y' \neq x$, $y' \in V(X; x)$. But then $x \in V(X; y') \subset V(X; y') \cup \{y'\} = V(X; y) \cup \{y\}$, and therefore $x \in V(X; y)$.

Now denote the vertices of X_* by w_a, $a \in A$. For each $a \in A$ choose a fixed element $x_a \in w_a$, and define $\sigma: X_* \to X$ by $\sigma w_a = x_a$. By the preceding paragraph σ is a monomorphism, proving (i).

→ Let $y \in x_*$. Then $V(X; x) \cup \{x\} = V(X; y) \cup \{y\}$. Hence either $x = y$ or $y \in V(X; x)$. Therefore $x_* \subset V(X; x) \cup \{x\}$. It follows that

$$V(X; x) \cup \{x\} = \bigcup \{z_* : z_* \in V(X_*; x_*) \cup \{x_*\}\}.$$

Hence $x_* \sim y_*$ implies $x \sim y$, i.e., $x_* = y_*$. Thus the natural projection $X_* \to X_{**}$ is an isomorphism. This proves (ii).

(iii) follows immediately from (ii). (iv) is obvious.

(6.5) LEMMA. *Let X be a graph. Then $R_0 = (X_*; C_{|x_*|} : x_* \in X_*)$ is a representation of X. It is minimal in the sense that if $R = (Y; C_{n_y} : y \in Y)$ is any representation of X, then X_* is isomorphic to a subgraph of Y.*

PROOF. It follows from the proof of (6.4) (ii) that for each $x \in X$ the maximal subgraph M_x of X with $V(M_x) = x_*$ is complete. This, together with the fact that $[x_*, y_*] \in X_*$ implies $[x, y] \in X$, proves the first part of the lemma. To prove the second part consider $X = \vee (C_{n_y}; Y)$, and define an equivalence $\sim_R$ on $V(X)$ by setting $(\alpha, y) \sim_R (\alpha', y')$ if and only if $y = y'$. Then it is immediate from (6.1) that for each $(\alpha, y) \in X$, $\{(\beta, z) \in X : (\beta, z) \sim_R (\alpha, y)\} \subset (\alpha, y)_*$. Denote the vertices of X_* by w_a, $a \in A$. For each class $w_a = (\alpha, y)_*$ choose a fixed $y_a \in Y$ such that $(\alpha_a, y_a) \in w_a$ for some α_a. Define $\sigma: X_* \to Y$ by $\sigma w_a = y_a$, $a \in A$. Then it is easily checked that σ is a monomorphism.

(6.6) COROLLARY. *A n.a.s.c. that $X \approx X_*$ is that all representations of X be trivial.*

(6.7) LEMMA. *Let $R = (Y; C_{n_y} : y \in Y)$ be a representation of X. Then $X_* \approx Y_*$.*

PROOF. Let $X = \vee (C_{n_y}; Y)$, $(\alpha, y) \in X$. Using the fact that

$$V(X; (\alpha, y)) \cup \{(\alpha, y)\} = \{(\beta, z) \in X : z \in V(Y; y) \cup \{y\}\},$$

one easily checks that

$$(\alpha, y)_* = \{(\alpha', y') \in X : y' \in y_*\}.$$

Hence $\tau: X_* \to Y_*$ given by $\tau((\alpha, y)_*) = y_*$, $(\alpha, y)_* \in X_*$, is an isomorphism.

(6.8) LEMMA. *Let $R = (Y; C_{n_y} : y \in Y)$ be a representation of X. If R and $R_0 = (X_*; C_{|x_*|} : x_* \in X_*)$ are isomorphic, then they are equivalent.*

PROOF. By (6.2) there is an isomorphism $\tau: X \approx V(C_{n_y}; Y)$. Since $\sim$ is invariant under isomorphisms it follows that $(\tau x)_* = \tau x_*$ for each $x \in X$. By the proof of (6.7), $\tau_*: X_* \approx Y_*$, where $\tau_* x_* = y_*$, and y is defined by $\tau x = (\alpha, y)$. $X_* \approx Y$ implies $Y_* \approx Y$. Hence by (6.4) (iii), τ_* can be continued to an isomorphism $\sigma: X_* \approx Y$ defined by $\sigma x_* = y$. Therefore $n_{\sigma x_*} = n_y = |(\alpha, y)_*| = |\tau x_*| = |x_*|$.

(6.9) LEMMA. *Let $\mathfrak{A} = \{A_e: e \in X\}$ be a complete cover of X. Then a n.a.s.c. that $X_* \approx X$ is that X contain no edge $e = [x, y]$ such that $V(X; x) - V(A_e) = V(X; y) - V(A_e)$. If $\mathfrak{A}$ is invariant, $e = [x, y] \in X$ is such that $x \sim y$, and if $e' \in E(X; x) \cup E(X; y)$, then $A_{e'} \supset A_e$ or $A_{e'} = e'$.*

PROOF. For any edge $e = [x, y] \in X$, $(V(X; x) \cup \{x\}) \cap V(A_e) = V(A_e)$.

To prove the second part of the lemma suppose that $e' = [x, y'] \notin A_e$ and $A_{e'} \neq e'$. Then there is a $z \in V(X; x)$ with $[y', z] \in A_{e'}$. $z \neq y$, otherwise $A_{e'} = A_e$. Let τ be the transposition (x, y). Since $x \sim y$, $\tau \in G(X)$. $[y', z] \in A_{e'} \cap A_{\tau e'}$, hence $A_{e'} = A_{\tau e'}$. Therefore $\tau e' = [y, y'] \in A_{e'}$, and hence $e \in A_{e'}$, a contradiction.

(6.10) COROLLARY. *If X is a quasi-derivative relative to the cover $\mathfrak{A}$, and if $x \in X$, there exists an $A \in \mathfrak{A}$ such that $x_* \subset V(A)$. This inclusion is proper if and only if X is not complete.*

PROOF. If $|x_*| = 1$ the statement is trivial, since by the maximality of $\mathfrak{A}$ either $|X| = 1$ or $|A| \geqq 2$ for each $A \in \mathfrak{A}$.

If $|x_*| \geqq 2$ let x, y be two distinct elements of x_*, e_0 the edge $[x, y]$. We claim that $x_* \subset V(A_{e_0})$. If not, let $z \in x_* - V(A_{e_0})$ and put $e_z = [x, z]$. $e_z \notin A_{e_0}$, hence by (6.9), $A_{e_z} = e_z$. Since A_{e_0} is complete, $V(A_{e_0}) \subset V(X; x) \cup \{x\}$ and therefore $V(A_{e_0}) \subset V(X; z)$. For $a \in A_{e_0}$ let $e_a = [a, z]$. $x \sim z$, $e_a \in E(X; z)$; hence by (6.9) either $A_{e_a} = e_a$ or $A_{e_a} = A_{e_z} = e_z$, but that is possible only when $a = z$. Hence $A_{e_a} = e_a$ for each $a \in A_{e_0}$. But then $B_0 = A_{e_0} \cup \bigcup \{(e_a): a \in A_{e_0}\}$ is complete, and $|B_0 \cap A_e| \leqq 1$ for each $e \notin B_0$. Put $B_e = B_0$ if $e \in B_0$, $B_e = A_e$ otherwise. Then $\mathfrak{B} = \{B_e: e \in X\}$ is a complete cover of X, and $\mathfrak{B} > \mathfrak{A}$, a contradiction. Hence $x_* \subset V(A_{e_0})$.

If X is complete, then $X = A_{e_0}$ and hence $x_* = V(A_{e_0})$.

If $x_* = V(A_{e_0})$ and X is not complete let $z \in V(X) - V(A_{e_0})$ be adjacent to some $a \in A_{e_0}$. Then $V(A_{e_0}) \subset V(X; z)$ and the argument of the preceding paragraph can be repeated, again leading to a contradiction.

(6.11) THEOREM. *Let X be a graph without isolated vertices. Then a n.a.s.c. that $(\partial X)_* \approx \partial X$ is that X have the following properties:* (i) *no vertex of X is adjacent to two distinct vertices of degree* 1; (ii) *no two vertices of degree* 2 *belong to a* 3*-circuit of X.*

PROOF. Necessity: Recall from (5.1) (ii) that a vertex e of ∂X is non-regular with respect to the natural cover $\mathfrak{E}$ if and only if e is incident with a vertex of degree 1 of X. Consequently, if e and e' are each incident with a vertex of degree 1, then $e \sim e'$. This proves the necessity of condition (i).

Let (x, y, z) be a 3-circuit of X such that $d(X; x) = d(X; y) = 2$. Put $[x, y] = e_x$, $[y, z] = e_y$, $[z, x] = e_z$. Then $\varepsilon = [e_y, e_z] \in E(\partial X)$, and $V(\partial X; e_y) -$

$V(E_e)=V(\partial X; e_x)-V(E_e)=\{e_x\}$. Hence by (6.9), $(\partial X)_*$ and ∂X are not isomorphic

Sufficiency: Without loss of generality we may assume that X is connected. If X is one of the exceptional graphs C_3, C_4, Q, ∂Q the sufficiency of conditions (i) and (ii) can easily be verified directly. Suppose, then, that X is none of the exceptional graphs, and that e_0, e_1 are two distinct vertices of ∂X with $e_0 \sim e_1$. Then $\varepsilon=[e_0, e_1]\in E(\partial X)$, i.e., $e_i=[x, x_i]$, $i=0, 1$, $x_0 \neq x_1$. If $\partial X=(\varepsilon)$, then $X=(e_0)\cup(e_1)$, and hence $d(X; x_i)=1$, $i=0, 1$, so that x violates hypothesis (i). Hence $\partial X \neq (\varepsilon)$, i.e., $A_i=E(\partial X; e_i)-\{\varepsilon\}\neq \square$, $i=0$ or 1. By (6.9) there is at most one $\varepsilon'\in A_i$ such that $\varepsilon'\notin E_e$; otherwise e_i would belong to more than two members of $\mathfrak{E}$ which is impossible. By condition (i) at least one of e_0 and e_1 is regular, say e_0. Hence there is exactly one $\varepsilon_0=[e_0, e_2]\in A_0$ with $E_{e_0}=\varepsilon_0$. Since $e_0\sim e_1$, $\varepsilon_1=[e_1, e_2]\in E(\partial X)$. But $E_{e_1}=\varepsilon_1$, otherwise $E_{e_1}=E_e$ and then $\varepsilon_0\in E_e$, a contradiction. Hence

$$E(\partial X; e_2)=\{\varepsilon_0, \varepsilon_1\}, \tag{1}$$

otherwise, e_2 belongs to more than two members of $\mathfrak{E}$. Therefore $x_i=e_i\cap e_2$, $i=0, 1$, and (x, x_0, x_1) form a 3-circuit of X. By (1), $d(X; x_i)=2$, $i=0, 1$, contrary to hypothesis (ii).

7. The derivative of graphs with multiple edges

The purpose of this section is to prove a theorem similar to (5.2) for graphs with multiple edges. The derivatives of such graphs are not quasi-derivatives, and while this prevents direct application of the results of section 4 the quotient operation introduced in section 6 makes it possible to reduce derivatives of graphs with multiple edges to derivatives of simple graphs.

In this section we shall mean by a graph a triple $X=(V(X), E(X), f_X)$, where $V(X)$ and $E(X)$ are any two disjoint sets, and f_X is a function of $E(X)$ into the set of unordered pairs of distinct elements of $V(X)$. If f_X is one-one X is called *simple*; any simple graph may be regarded as a graph in the sense of section 1. We shall say two edges $e_1, e_2\in E(X)$ are *adjacent* if and only if $e_1\neq e_2$ and the two pairs $f_X e_1$ and $f_X e_2$ have a common element. We define the *derivative* of X exactly as in section 1, and note that ∂X is a simple graph [7]).

It is clear that every graph contains a maximal simple subgraph, and that any two maximal simple subgraphs of X are isomorphic.

(7.1) LEMMA. *Let X be a graph, Y a maximal simple subgraph of X. For $e\in E(Y)$ let $m_e=|f_X^{-1}f_Y e|$. Then $R=(\partial Y; C_{m_e}: e\in\partial Y)$ is a representation of ∂X.*

PROOF. Put $\vee(C_{m_e}; \partial Y)=Z$. For each $e\in\partial Y$ let τ_e be a one-one function of $V(C_{m_e})$ onto $f_X^{-1}f_Y e$. Define $\varrho: Z\to\partial X$ by $\varrho(\alpha, e)=\tau_e\alpha$, $\alpha\in C_{m_e}$, $e\in\partial Y$. Clearly ϱ is onto. By definition $\tau_e\alpha\in f_X^{-1}f_Y e$. Hence if $\tau_e\alpha=\tau_{e'}\alpha'$, then $f_X^{-1}f_Y e=f_X^{-1}f_Y e'$, and therefore $e=e'$. Thus ϱ is one-one. Let $[(\alpha, e), (\alpha', e')]\in Z$. If $e=e'$, then $[\alpha, \alpha']\in C_{m_e}$, i.e., $\alpha\neq\alpha'$. Hence $\tau_e\alpha\neq\tau_e\alpha'$, and since both

[7]) This differs from the definition given in [2].

edges belong to $f_X^{-1}f_Y e$, $[\tau_e\alpha, \tau_e\alpha'] \in \partial X$. If $e \neq e'$, then $[e, e'] \in \partial Y$, hence again $[\tau_e\alpha, \tau_{e'}\alpha'] \in \partial X$. It follows that ϱ is a homomorphism. The proof that ϱ^{-1} is a homomorphism is similar.

(7.2) COROLLARY. *Let X be a graph, Y a maximal simple subgraph of X. Then $(\partial X)_* \approx (\partial Y)_*$.*

PROOF. (7.1), (6.7).

(7.3) LEMMA. *Let Y be a simple graph, $R = (\partial Y; C_{n_e}: e \in \partial Y)$ a representation of X. Then there is a graph X_R such that* (i) $\partial X_R \approx X$, (ii) *Y is a maximal simple subgraph of X_R, and* (iii) *for each $e \in \partial Y$, $|f_{X_R}^{-1} f_Y e| = n_e$. If $R = (\partial Y; C_{n_e}: e \in \partial Y)$ and $S = (\partial Z; C_{m_e}: e \in \partial Z)$ are two representations of X with Y and Z simple, then a n.a.s.c. that $X_R \approx X_S$ is that R and S be equivalent.*

PROOF. For each $e \in E(Y)$ let A_e be a set disjoint from $V(Y)$ such that (1) $e \in A_e$, (2) $|A_e| = n_e$, and (3) $A_e \cap A_{e'} = \square$ if $e \neq e'$. Define X_R by $V(X_R) = V(Y)$, $E(X_R) = \bigcup \{A_e: e \in E(Y)\}$, $f_{X_R} a = f_Y e$, where $a \in A_e$, $e \in E(Y)$. Then (i), (ii), (iii) are immediate from the construction of X_R.

As for the second part of the lemma, sufficiency is obvious. To prove necessity assume $\eta: X_R \approx X_S$. Then $\eta: Y \approx Z$ and $n_e = |f_{X_R}^{-1} f_Y e| = |f_{X_S}^{-1} f_Y \eta e| = m_{\eta e}$, $e \in E(Y)$. Thus R and S are equivalent.

It follows from the preceding sequence of lemmas that the problem of determining for a given graph X the collection of all non-isomorphic graphs X_α with $\partial X_\alpha \approx \partial X$ is equivalent with the problem of determining all inequivalent representations $R_\alpha = (\partial Z_\alpha; C_{n_e}: e \in \partial Z_\alpha)$ of ∂X for which Z_α is simple.

In order to solve this problem under certain special assumptions we make the following observations. Let Y be a connected simple graph. It is clear from (6.10) that $(\partial Y)_*$ is a derivative with the property that for each $\varepsilon \in E((\partial Y)_*)$, E_ε has at most one non-regular vertex. Moreover, *if $(\partial Y)_*$ is regular*, (i.e., if every vertex of $(\partial Y)_*$ is regular) *then $(\partial Y)_* \approx \partial Y$.* (The converse of this is false.) Also from (6.11) it follows that *$(\partial Y)_*$ is regular if and only if* (i) *Y has no vertex of degree 1, and* (ii) *Y has no 3-circuit containing two vertices of degree 2.* We can now prove the following theorem.

(7.4) THEOREM. *Let X, X' be connected graphs such that* (i) *$(\partial X)_*$ is regular, and* (ii) *$\partial X \approx \partial X'$. Then $X \approx X'$.*

PROOF. Let Y, Y' be maximal simple subgraphs of X and X', respectively. (ii) implies $(\partial X)_* \approx (\partial X')_*$, hence by (7.2), $(\partial Y)_* \approx (\partial Y')_*$. By (i), $(\partial Y)_*$ is regular, hence $(\partial Y)_* \approx \partial Y$. Also $(\partial Y')_*$ must be regular, and therefore $(\partial Y')_* \approx \partial Y'$. Hence $\partial Y \approx \partial Y'$. Since ∂Y and $\partial Y'$ are regular, neither Y nor Y' is isomorphic to C_3. Hence by (5.2), $Y \approx Y'$. Let $\eta: Y \approx Y'$. Then by (6.8), $m_e = |f_X^{-1} f_Y e| = |f_{X'}^{-1} f_{Y'} \eta e| = m'_{\eta e}$ for each $e \in E(Y)$. Hence $X \approx X'$.

(7.5) COROLLARY. *Let X, X' be cyclically connected graphs such that* (i) *the maximal simple subgraphs of X and X' are not isomorphic to C_3, and* (ii) *$\partial X \approx \partial X'$. Then $X \approx X'$.*

(7.6) The other case which we shall consider is that of a simple connected graph X not isomorphic to C_3. We shall only sketch the result leaving the argument to the reader.

Let M be the set of all vertices of X which are adjacent to some vertex of degree 1. For $x \in M$ let W_x be any subset of $V(X; x)$ such that $d(X; y) = 1$ for eacl $v \in W_x$; for $x \notin M$ put $W_x = V(X; x)$. To each $y \in W_x$ assign a subset $E_{xy} \subset E($ $x)$ such that (i) $\cup \{E_{xy}: y \in W_x\} = \{[x, z] \in X: d(X; z) = 1\}$, and (ii) $E_{xy} \cap E_{xz} = \square$ if $y \neq z$, $y, z \in W_x$.

By $\{C_\alpha: \alpha \in A\}$ denote the set of all 3-circuits of X which contain two vertices of degree 2. Let $C_\alpha = (x_\alpha, y_\alpha, z_\alpha)$ and suppose that $d(X; x_\alpha) = d(X; y_\alpha) = 2$. Then z_α necessarily is a cut vertex of X. Note that it is possible that $z_\alpha = z_\beta$ for $\alpha \neq \beta$.

Let $B \subset A$ be given. Define Y by

$$V(Y) = \cup \{W_x: x \in X\},$$

$$E(Y) = E(X),$$

$$f_Y e = \begin{cases} e, & \text{if } e \notin \cup \{E_{xy}: y \in W_x,\ x \in X\} \cup \cup \{C_\alpha: \alpha \in B\} \\ [x, y], & \text{if } e \in E_{xy} \\ e, & \text{if } e = [x_\alpha, y_\alpha],\ \alpha \in B \\ [y_\alpha, z_\alpha], & \text{if } e = [x_\alpha, z_\alpha], \text{ or } e = [y_\alpha, z_\alpha],\ \alpha \in B. \end{cases}$$

Then $\partial Y \approx \partial X$. Conversely, any graph Y with $\partial Y \approx \partial X$ can be obtained by this construction (i.e., by a choice of sets W_x, E_{xy}, and B). Note, however, that the graphs arising from different choices of these sets are not necessarily non-isomorphic.

References

[1] Izbicki, H.: Über die explizite Automorphismengruppe von speziellen Graphen. J. reine angew. Math. **203**, 35–39 (1960).

[2] — Graphentransformationen. Mh. Math. Phys. **64**, 135–175 (1960).

[3] Sabidussi, G.: The composition of graphs. Duke Math. J. **26**, 693–696 (1959).

[4] Whitney, H.: Congruent graphs and the connectivity of graphs. Amer. J. Math. **54**, 150–168 (1932).

Tulane University and McMaster University, Hamilton, Ontario (Canada)

(Eingegangen am 23. November 1960)

SCHMIDT, H.
Math. Zeitschr. 76, 402/403 (1961)

Zur Charakterisierung der zyklischen Funktionen durch eine Funktionalgleichung

Von
HERMANN SCHMIDT

Wie wir vor kurzem hervorgehoben haben[1]), genügt das Additionstheorem

$$(1) \qquad f_\nu(x+y) = \sum_{j=1}^{n} f_j(x)\, f_{\nu-j+1}(y) \qquad (\nu = 1, 2, \ldots, n)$$

der sog. „zyklischen" Funktionen keineswegs zu deren Charakterisierung; vielmehr sind in der Lösung n Konstante willkürlich. Wohl aber genügt für $f_1(x)$ hierzu das Gegenstück zu der öfters und unter sehr allgemeinen Annahmen[2]) untersuchten Funktionalgleichung des Cosinus

$$f(x+y) + f(x-y) = 2f(x)\, f(y).$$

Dies mag im folgenden kurz ausgeführt werden, wobei wir freilich wegen des unvermeidlichen Auftretens komplexer Argumente nicht ohne funktionentheoretische Voraussetzungen auskommen.

Für die zyklischen Funktionen folgt zunächst aus (1) und mit Rücksicht auf die Reihendarstellungen [die Potenzexponenten für $f_\nu(x)$ sind kongruent $\nu - 1 \bmod n$]

$$\sum_{\varkappa=1}^{n} f_\nu(x + \varepsilon^{\varkappa-1} y) = \sum_{\varkappa,j=1}^{n} f_j(x)\, f_{\nu-j+1}(\varepsilon^{\varkappa-1} y) = \sum_{\varkappa,j=1}^{n} f_j(x)\, \varepsilon^{(\nu-j)(\varkappa-1)}\, f_{\nu-j+1}(y)$$

$$= n\, f_\nu(x)\, f_1(y)$$

und für $\nu = 1$ insbesondere

$$\sum_{\varkappa=1}^{n} f_1(x + \varepsilon^{\varkappa-1} y) = n\, f_1(x)\, f_1(y).$$

Umgekehrt gilt nun der

SATZ. *Es sei n eine natürliche Zahl $\geqq 3$, $\varepsilon = e^{\frac{2\pi i}{n}}$ und $\mathfrak{S}$ das Sterngebiet $\{\varepsilon^\lambda x + \varepsilon^\mu y\}$, wo $0 \leqq \lambda, \mu \leqq n-1$; $x, y \in \mathfrak{i} = \langle 0, r)$. Befriedigt dann die in $\mathfrak{S}$ reguläre Funktion $f(z)$ $(\not\equiv 0)$ für $x, y \in \mathfrak{i}$ die Funktionalgleichung*

$$(2) \qquad \sum_{\varkappa=1}^{n} f(x + \varepsilon^{\varkappa-1} y) = n\, f(x)\, f(y),$$

so ist $f(z) = f_1(\gamma z)$, wo γ eine Konstante bedeutet.

---

[1]) HERM. SCHMIDT (mit einem Zusatz von O. E. GHEORGHIU), diese Z. **76**, 46–50 (1961), insbes. Satz 1.

[2]) Zuletzt von S. KUREPA [Canad. J. Math. **12**, 45–50 (1960)] für von einem reellen Parameter x abhängige stetige Abbildungen eines Hilbertschen Raumes in sich.

BEWEIS. Für $x=y=0$ ergibt sich $f(0)=f(0)^2$, also

(3) $\qquad f(0)=1,$ da bei $f(0)=0$ aus (2) für $y=0$ folgte $f(x)\equiv 0$.

Für $x=0$, $y\in\mathfrak{i}$ liefert jetzt (2) die (von selbst in $\mathfrak{S}$ gültige) Beziehung

$$\sum_{\varkappa=1}^{n} f(\varepsilon^{\varkappa-1}y)=n\,f(y).$$

Ersetzt man y durch $\varepsilon y\,(\in\mathfrak{S})$, so folgt der Reihe nach $f(y)=f(\varepsilon y)=\cdots=f(\varepsilon^{n-1}y)$, woraus durch Differentiation

(4) $\qquad f^{(\lambda)}(y)=\varepsilon^{\nu\lambda}f^{(\lambda)}(\varepsilon^{\nu}y) \qquad (\nu=0,1,\ldots,n-1;\ \lambda=1,2,\ldots).$

Endlich ergibt λ-maliges Differenzieren von (2) nach x ($x,\ y\in\mathfrak{i}$)

$$\sum_{\varkappa=1}^{n} f^{(\lambda)}(x+\varepsilon^{\varkappa-1}y)=n\,f^{(\lambda)}(x)\,f(y),$$

woraus für $x=0$ vermöge (4)

$$n\,f^{(\lambda)}(0)\,f(y)=\sum_{\varkappa=1}^{n}\varepsilon^{-(\varkappa-1)\lambda}f^{(\lambda)}(y), \quad =0 \quad \text{für} \quad \lambda=1,2,\ldots,n-1,$$

somit

(5) $\qquad f^{(\lambda)}(0)=0 \quad (\lambda=1,2,\ldots,n-1).$

Dagegen kommt für $\lambda=n$

$$f^{(n)}(y)=f^{(n)}(0)\,f(y),$$

somit mit Rücksicht auf (3), (5) $f(y)=f_1(\gamma y)$ mit $\gamma^n=f^{(n)}(0)$, was dann identisch in $\mathfrak{S}$ gilt.

BEMERKUNG. Natürlich bleibt die Betrachtungsweise auch für $n=2,1$ anwendbar, wobei es genügt, für $f(x)$ zwei- bzw. einmalige Differenzierbarkeit in einem passenden reellen Intervall anzunehmen; doch hat man ja in diesen Fällen viel allgemeinere Ergebnisse.

Würzburg, Mathematisches Institut der Universität

(Eingegangen am 13. April 1961)

CARTWRIGHT, M. L., and E. F. COLLINGWOOD
Math. Zeitschr. 76, 404–410 (1961)

The radial limits of functions meromorphic in a circular disc

By
M. L. CARTWRIGHT and E. F. COLLINGWOOD

1. Let $w=f(z)$ be meromorphic in the unit disc $|z|<1$, denoted by U, which it maps on the w-plane or, by stereographic projeciton, on the Riemann sphere S, and let $\varkappa$ denote the circumference $|z|=1$. We denote by $D(f)$ the set of points $e^{i\theta}\in\varkappa$ at which the radial limit $\lim_{r\to 1} f(re^{i\theta})=\Gamma_\varrho(f, e^{i\theta})$ exists and by $\Gamma_\varrho(f)$ the set $\bigcup_{e^{i\theta}\in\varkappa}\Gamma_\varrho(f, e^{i\theta})$ of all the radial limits of $f(z)$. For a set $A<\varkappa$ the set $\bigcup_{e^{i\theta}\in A}\Gamma_\varrho(f, e^{i\theta})$ is denoted by $\Gamma_\varrho(f, A)$.

It was proved by PRIVALOV[1]) that

If the set $D(f)\cap A$ is both of category II[2]) *and metrically dense*[3]) *on an arc $A<\varkappa$ then, either $\Gamma_\varrho(f)$ contains a closed set of positive capacity, or $f(z)$ is constant.*

This theorem may now be improved in two ways; the hypothesis may be weakened by dropping the metrical condition and the conclusion strengthened. As a special case of a more general theorem we shall prove

THEOREM 1. *If $f(z)$ is meromorphic in U and if $D(f)\cap A$ is of category* II *on an arc $A<\varkappa$ then, either $\Gamma_\varrho(f, A)$ is of positive linear measure, or $f(z)$ is constant.*

2. Our approach is that of the theory of cluster sets. In the usual notation[4]) $C(f, e^{i\theta})$ is the *complete cluster set* of $f(z)$ at $e^{i\theta}$ defined as follows: $\alpha\in C(f, e^{i\theta})$ if there is a sequence $\{z_n\}<U$ such that $\lim_{n\to\infty} z_n=e^{i\theta}$ and $\lim_{n\to\infty} f(z_n)=\alpha$. The set $C(f, e^{i\theta})$ is closed; and if its complement $\mathfrak{C}C(f, e^{i\theta})$ is empty we say that $C(f, e^{i\theta})$ is *total*, otherwise it is *sub-total*. Similarly, the radial cluster set $C_\varrho(f, e^{i\theta})$ is defined by imposing the further condition $z_n=r_n e^{i\theta}$, where r_n is real and $0<r_n<1$.

A point $e^{i\theta}\in\varkappa$ such that $C(f, e^{i\theta})$ is total we call a *Weierstrass point* of $f(z)$ and denote the set of all such points on $\varkappa$ by $W(f)$ and the sub-set of them on $A<\varkappa$ by $W(f, A)$. A point $e^{i\theta}\in\varkappa$ such that $f(z)$ has a unique finite or infinite limit, denoted by $f(e^{i\theta})$, as $z\to e^{i\theta}$ in every Stolz angle $\Delta<U$ having $e^{i\theta}$ for its vertex is a *Fatou point* of $f(z)$. The set of all Fatou points of $f(z)$

[1]) PRIVALOV [8]. The proof is reproduced in [9] pp. 231–232. The numbers in square brackets refer to the list of references at the end.

[2]) We write category I and II for Baire's first and second category.

[3]) A set is metrically dense on an interval if the measure of its intersection with any sub-interval is positive.

[4]) The general reference for notation and definitions is to [5] and [1].

is denoted by $F(f)$ and the corresponding set $\bigcup_{e^{i\phi}\in F(f)} f(e^{i\phi})$ by $\Gamma_F(f)$. We write $F(f, A)$ for $F(f)\cap A$ and $\Gamma_F(f, A)$ for the corresponding set of angular limits on A.

An arc of $\varkappa$ in some neighbourhood of which either $f(z)$ or, for some $a\neq\infty$, $1/|f(z)-a|$ is uniformly bounded is called a *Fatou arc* for $f(z)$. It follows from the well-known theorem of Fatou that *almost all points of a Fatou arc are Fatou points of* $f(z)$.

A *pointwise asymptotic value of* $f(z)$ is a value α such that, for some $e^{i\phi}\in\varkappa$ and some Jordan arc $\gamma<U$ terminating at $e^{i\phi}$, $f(z)\to\alpha$ as $z\to e^{i\phi}$ along γ. The set of all such values is denoted by $\Gamma_P(f)$ and the subset of them associated with points $e^{i\phi}\in A<\varkappa$ by $\Gamma_P(f, A)$.

We note that if A, which may be the whole of $\varkappa$, is a Fatou arc, then, by Lindelöf's theorem on asymptotic values,

$$\Gamma_F(f, A)=\Gamma_Q(f, A)=\Gamma_P(f, A), \tag{1}$$

while for any set $A<\varkappa$,

$$\Gamma_F(f, A)<\Gamma_Q(f, A)<\Gamma_P(f, A). \tag{2}$$

3. Suppose now that $e^{i\phi}\in\mathfrak{C}W(f)=\varkappa-W(f)$. Then $\mathfrak{C}C(f, e^{i\phi})$ is not empty and is thus open and contains a circular disc of centre a, say, and radius $\varepsilon>0$. There is therefore a circular disc d of centre $e^{i\phi}$ and radius $\delta>0$ such that, in $d\cap U$, $f(z)$ does not meet the disc of centre a and radius $\varepsilon/2$. The arc $\varkappa\cap d$ is therefore an open Fatou arc of $f(z)$. We have thus proved

LEMMA 1. *If an arc A of $\varkappa$ contains no Fatou arc of $f(z)$, then $A<W(f)$.*

We also require the following lemma which we proved in a previous paper[5]).

LEMMA 2. *If B is a Fatou arc of $f(z)\neq$ constant, then for any pair ϑ_1, ϑ_2 such that $e^{i\vartheta_1}\in B\cap F(f)$, $e^{i\vartheta_2}\in B\cap F(f)$ and $f(e^{i\vartheta_1})\neq f(e^{i\vartheta_2})$, the projection on the straight line L joining $f(e^{i\vartheta_1})$ and $f(e^{i\vartheta_2})$ of the set $\Gamma_F(f, B)$ contains L; and hence $\Gamma_F(f, B)$ is of positive linear measure.*

By the theorem of F. and M. Riesz $e^{i\vartheta_1}$ and $e^{i\vartheta_2}$ can always be found satisfying the conditions of the lemma and arbitrarily near to any given pair of points of B.

4. We shall now prove our main theorem from which Theorem 1 follows.

We denote by $E(f)$ the set of points $e^{i\phi}\in\varkappa$ such that $C_Q(f, e^{i\phi})$ is sub-total. By a theorem of COLLINGWOOD[6]), valid for continuous functions,

$$C_Q(f, e^{i\phi})=C(f, e^{i\phi}) \tag{3}$$

except perhaps in a set of points $e^{i\phi}$ of category I *on* $\varkappa$. Hence, if on an arc $A<\varkappa$ the set $E(f)\cap A$ is of category II the set $\mathfrak{C}W(f)\cap A$ is also of category II and therefore not empty. By Lemma 1. A therefore contains a Fatou arc B

[5]) [5], Lemma 4. This is proved for a mapping on the plane but the conclusion on $\Gamma_F(f, B)$ is plainly valid also for a mapping on S.

[6]) [3], Theorem 1. See also [2].

and it follows from Lemma 2. that, if $f(z)$ is not constant, $\Gamma_\varrho(f, A) > \Gamma_\varrho(f, B) = \Gamma_F(f, B)$ is of positive linear measure. We have thus proved

THEOREM 2. *If $f(z)$ is meromorphic in U and $E(f)$ is of category* II *on some arc $A < \varkappa$, then either $\Gamma_\varrho(f, A)$ is of positive linear measure or $f(z)$ is constant.*

This theorem contains Theorem 1 since $D(f) < E(f)$.

5. We write $\mathrm{m}\Gamma$ for the linear measure of the set Γ and $\operatorname{cap}\Gamma$ for its logarithmic capacity.

In the usual notation $R(f, e^{i\vartheta})$ is the set of values α such that $f(z_n) = \alpha$ for all z_n belonging to a sequence $\{z_n\} < U$, $\lim_{n\to\infty} z_n = e^{i\vartheta}$. The complement $S - R(f, e^{i\vartheta})$ is denoted by $\mathfrak{C}R(f, e^{i\vartheta})$. A point $e^{i\vartheta}$ for which $\operatorname{cap}\mathfrak{C}R(f, e^{i\vartheta}) = 0$ is termed a *Frostman point* and the set of such points of $\varkappa$ is denoted by $H(f)$[7]). A point for which $\mathfrak{C}R(f, e^{i\vartheta})$ contains at most two points is a *Picard point* and the set of Picard points is denoted by $P(f)$. The notion of a Picard point may be generalised in the following way. Denote by $P_n(f)$, $0 \leqq n < \infty$, the set of points $e^{i\vartheta} \in \varkappa$ such that $\mathfrak{C}R(f, e^{i\vartheta})$ contains at most n values, and by $P_{\aleph_0}(f)$ the set for which $\mathfrak{C}R(f, e^{i\vartheta})$ is at most enumerable. Then, plainly, for any finite $n \geqq 0$ we have $P_n(f) < P_{n+1}(f) < P_{\aleph_0}(f)$ and, in particular,

$$P(f) = P_2(f) > P_0(f) \cup P_1(f).$$

We shall now prove

THEOREM 3. *If $f(z)$ is meromorphic in U and if*

(i) *for some finite $n \geqq 0$ the set $\Gamma_F(f)$ contains at most n values, then*

$$\varkappa < P(f) \cup P_n(f);$$

while if

(ii) *$\Gamma_F(f)$ is at most enumerable, then*

$$\varkappa < P_{\aleph_0}(f).$$

To prove this we rely on theorems proved in our previous paper [5]. The following definitions and notation are required. By $\Gamma_P(f, |\vartheta' - \vartheta| < \eta)$ and $\Gamma_F(f, |\vartheta' - \vartheta| < \eta)$ we denote the sub-sets of $\Gamma_P(f)$ and $\Gamma_F(f)$ associated with points of the arc $(e^{i(\vartheta-\eta)}, e^{i(\vartheta+\eta)})$ of $\varkappa$ which contains the point $e^{i\vartheta}$. For the intersections of these sets for indefinitely decreasing η we write

$$\chi_P(f, e^{i\vartheta}) = \bigcap_\eta \Gamma_P(f, |\vartheta' - \vartheta| < \eta)$$

and

$$\chi_F(f, e^{i\vartheta}) = \bigcap_\eta \Gamma_F(f, |\vartheta' - \vartheta| < \eta).$$

The set $\Gamma_P(f)$ is a subset of $\Gamma(f)$, the set of all the asymptotic values of $f(z)$, which may contain values which are not pointwise, i.e. limits of $f(z)$ on paths in U converging not to end-points on $\varkappa$ but to arcs of $\varkappa$ or to the whole of $\varkappa$. We denote by $\Gamma(f, |\vartheta' - \vartheta| < \eta)$ the sub-set of $\Gamma(f)$ associated with asymptotic paths which converge to arcs or end-points contained in the

[7]) [5], p. 138.

arc $(e^{i(\vartheta-\eta)}, e^{i(\vartheta+\eta)})$ of $\varkappa$, and we write

$$\chi(f, e^{i\vartheta}) = \bigcap_{\eta} \Gamma\big(f, |\vartheta' - \vartheta| < \eta\big).$$

To prove the theorem suppose that $e^{i\vartheta}$ is not a point of $P(f)$. We can therefore find $\eta > 0$ such that in the sub-domain $V(\vartheta)$ of U bounded by the arc $(e^{i(\vartheta-\eta)}, e^{i(\vartheta+\eta)})$ containing $e^{i\vartheta}$ and the circular arc of centre $e^{i\vartheta}$ passing through the points $e^{i(\vartheta-\eta)}$ and $e^{i(\vartheta+\eta)}$ the function $f(z)$ omits at least three values. Under a conformal mapping of $V(\vartheta)$ onto a circular disc the sets $\Gamma\big(f, |\vartheta' - \vartheta| < \eta\big)$ and $\Gamma_P\big(f, |\vartheta' - \vartheta| < \eta\big)$ are invariant and it follows ([5], Theorem 6) that

$$\Gamma\big(f, |\vartheta' - \vartheta| < \eta\big) = \Gamma_P\big(f, |\vartheta' - \vartheta| < \eta\big) = \Gamma_F\big(f, |\vartheta' - \vartheta| < \eta\big)$$

and hence that

$$\chi(f, e^{i\vartheta}) = \chi_F(f, e^{i\vartheta}). \tag{4}$$

Now, by hypothesis, $\Gamma_F\big(f, |\vartheta' - \vartheta| < \eta\big) \subset \Gamma_F(f)$ is either finite or at most enumerable and thus of linear measure zero. Therefore ([5], Theorem 16 (ii), (19.3))

$$\mathscr{C}R(f, e^{i\vartheta}) \subset \chi(f, e^{i\vartheta})$$

and so, in virtue of (4),

$$\mathscr{C}R(f, e^{i\vartheta}) \subset \chi_F(f, e^{i\vartheta}).$$

But $\chi_F(f, e^{i\vartheta}) \subset \Gamma_F(f)$ and, since $e^{i\vartheta} \in \varkappa - P(f)$, both parts of the theorem follow.

A counter-example shows that Theorem 3 is best possible in the sense that the set $P(f)$ cannot be eliminated in (i) even if $n = 0$. We recall ([5], p. 94) the definition of the set $\Gamma_S(f)$ of *spirally asymptotic values* of $f(z)$. $\alpha \in \Gamma_S(f)$ if there is a continuous path $z = z(t)$, $0 < t < 1$, such that $\lim_{t\to 1} |z(t)| = 1$, $|\arg z(t)|$ is unbounded and $\lim_{t\to 1} f\big(z(t)\big) = \alpha$. The true spirals for which $\lim_{t\to 1} \arg z(t) = \pm\infty$ give a subset of $\Gamma_S(f)$. We can construct a function $f(z)$ regular and $\neq 0$ in U such that $(0) \cup (\infty) \subset \overline{\Gamma}_S(f)$, so that $\Gamma_F(f)$ must be empty while $\mathscr{C}R(f, e^{i\vartheta})$ contains both 0 and ∞ for every $e^{i\vartheta} \in \varkappa$. For the construction we refer to [4a], §2. The modular function $\mu(z)$, for which 0, 1, ∞ are the sole values belonging to $\mathscr{C}R(\mu)$ and to $\Gamma_F(\mu)$,and the function $\log \mu(z) = Q(z)$, for which $\mathscr{C}R(Q)$ and $\Gamma_F(Q)$ both consist of the values ∞, $\pm 2n\pi i$, $n = 0, 1, 2, \ldots$, are familiar examples which show that the theorem is also best possible when $\Gamma_F(f)$ contains more than two values.

6. The results now available, brought together, lead to a simple scheme of relations between the metrical or arithmetical properties of $\Gamma_F(f)$ and the nature of $\varkappa$ as a singular boundary. We set this out as

THEOREM 4. *Let $f(z)$ be meromorphic in U. Then*

(i) $\mathrm{m}\,\Gamma_F(f) = 0 \Rightarrow \varkappa \subset W(f)$;

(ii) $\mathrm{cap}\,\Gamma_F(f) = 0 \Rightarrow \varkappa \subset H(f)$;

(iii) *$\Gamma_F(f)$ enumerable* $\Rightarrow \varkappa \subset P_{\aleph_0}(f)$;

(iv) $\Gamma_F(f)$ *contains at most* n *values*

$$\Rightarrow \begin{cases} \varkappa < P_n(f) & \text{if } 2 < n < \infty \\ \varkappa < P(f) & \text{if } 0 \leqq n \leqq 2. \end{cases}$$

The implication (i) is an immediate consequence of Lemmas 1 and 2. (ii) is a known result[8]), but we give an alternative proof based on another theorem of PRIVALOV[9]), namely the implication

(v) $\operatorname{cap} \Gamma_F(f) = 0 \Rightarrow \mathrm{m}\, F(f) = 0.$

Now $\mathrm{m}F(f)=0$ implies that the set $F_1(f)$ of points $e^{i\theta} \in \varkappa$ at which the set $F(f)$ is of density 1 is empty. By a theorem of LITTLEWOOD[10]) every point of $\varkappa$ belongs either to $H(f)$ or to $F_1(f)$ so that (ii) follows from (v). In (iii) and (iv) we have re-stated Theorem 3.

COROLLARY 1. *If* $f(z)$ *is meromorphic in* U *and possesses at least one spirally asymptotic value, then* $\varkappa < P(f)$.

For, if $\Gamma_S(f)$ contains one value only, then $\Gamma_F(f)$ is either empty or contains this same value and no other; or, if $\Gamma_S(f)$ contains more than one value $\Gamma_F(f)$ is empty.

This corollary supplements some recent theorems of SEIDEL on functions having spirally asymptotic values[11]).

To prove our next corollary we require

LEMMA 3. *If* $f(z)$ *is meromorphic in* U, *then*

$$F(f') < F(f),$$

where $f'(z)$ *is the derivative of* $f(z)$.

Let $e^{i\theta} \in F(f')$ and $\Delta(\vartheta)$ be any Stolz angle in U having $e^{i\theta}$ as its vertex. Denote by $\delta(\varrho, \vartheta)$ the disc $|z - e^{i\theta}| < \varrho$ and by $\Delta(\varrho, \vartheta)$ the intersection $\Delta(\vartheta) \cap \delta(\varrho, \vartheta)$. Now let $\alpha = f'(e^{i\theta}) = \lim_{r \to 1} f'(r e^{i\theta})$. Then, given $\varepsilon_n > 0$ we can find ϱ_n such that $|f'(z) - \alpha| < \varepsilon_n$ in. $\Delta(\varrho_n, \vartheta)$. Let z, z_0 be two points in $\Delta(\varrho_n, \vartheta)$. Then

$$\begin{aligned} f(z) &= f(z_0) + \int_{z_0}^{z} f'(z)\, dz \\ &\doteq f(z_0) + \alpha \int_{z_0}^{z} dz + \int_{z_0}^{z} (f'(z) - \alpha)\, dz \\ &= f(z_0) + \alpha (z - z_0) + \varphi(z, z_0), \end{aligned}$$

where $|\varphi(z, z_0)| < \varepsilon_n |z - z_0| < 2\varepsilon_n \varrho_n$, and so

$$|f(z) - f(z_0)| < 2\varepsilon_n (\alpha + \varrho_n).$$

---

[8]) [6], p. 100.

[9]) [8], The proof is reproduced in [9] pp. 210–212. The theorem, which generalises a previous theorem of FROSTMAN and R. NEVANLINNA [7], p. 209, was re-discovered by TSUJI [10], who gave a different proof.

[10]) [5], Theorem 21 (p. 138).

[11]) [13].

Denote the cluster set of $f(z)$ in the Stolz angle $\Delta(\vartheta)$ in the usual way by $C_{\Delta(\vartheta)}(f, e^{i\vartheta})$ and let $\beta \in C_{\Delta(\vartheta)}(f, e^{i\vartheta})$. Then, given $\eta > 0$, we can find a sequence $\{z_n\} < \Delta(\vartheta)$, $|z_n| < |z_{n+1}|$, $\lim_{n\to\infty} z_n = e^{i\vartheta}$, such that $|f(z_n) - \beta| < \eta$. Suppose, as we may, that $|e^{i\vartheta} - z_n| < \varrho_n$ so that for $z \in \Delta(\varrho_n, \vartheta)$ we have

$$\begin{aligned} f(z) &= f(z_n) + \alpha(z - z_n) + \varphi(z, z_n) \\ &= \beta + (f(z_n) - \beta) + \alpha(z - z_n) + \varphi(z, z_n) \end{aligned}$$

and hence

$$|f(z) - \beta| < |f(z_n) - \beta| + 2\varepsilon_n(\alpha + \varrho_n) < 2\eta$$

by choice of n. But η is arbitrary and it follows that $C_{\Delta(\vartheta)}(f, e^{i\vartheta})$ is of zero diameter and therefore consists only of the value β. The convergence of $f(z)$ to β is uniform in $\Delta(\vartheta)$ so that $e^{i\vartheta} \in F(f)$ and the lemma is proved.

Lemma 3 at once gives

COROLLARY 2. *Let $f(z)$ be meromorphic in U and $f^{(\nu)}(z)$ its ν'th derivative. Then, for any ν,*

(i) $\mathrm{m}\,\Gamma_F(f) = 0 \Rightarrow \varkappa < W(f^{(\nu)})$;

(ii) $\mathrm{cap}\,\Gamma_F(f) = 0 \Rightarrow \varkappa < H(f^{(\nu)})$;

(iii) $\Gamma_F(f)$ *enumerable* $\Rightarrow \varkappa < P_{\aleph_0}(f^{(\nu)})$;

(iv) $\Gamma_F(f)$ *contains at most n values*

$$\Rightarrow \begin{cases} P_n(f^{(\nu)}) & \text{if } 2 < n < \infty \\ P(f^{(\nu)}) & \text{if } 0 \leqq n \leqq 2. \end{cases}$$

For $\Gamma_F(f^{(\nu)}) < \Gamma_F(f)$.

6. We conclude with an observation on the content of Lemma 2 and Theorem 1. As regards Lemma 2, the question may be asked whether $\Gamma_F(f, B)$ necessarily contains a continuum? The answer is in the negative even in the case where every point of the Fatou arc B belongs to $F(f)$. As counter example we cite a schlicht function $f(z)$ which maps U onto a bounded domain D, constructed by CARATHÉODORY[12]), whose prime ends are all of the first or second kinds (each thus having an accessible point, so that $\varkappa < F(f)$) and such that the accessible points of the boundary form a set which is rendered totally disconnected by the removal of a single (dispersion) point, p say. Such a set contains no continuum. For, any continuum contained in the set must contain p; and, by a well-known theorem[13]), any such continuum contains a sub-continuum which does not contain p. We do not repeat the construction of this domain D since it is given in detail by CARATHÉODORY in sections 49 and 51 of his paper. It is enough to say that in that construction the origin O is the impression of all the prime ends of D of the first kind and is not the accessible point of any of the non-enumerable set of prime

[12]) [11], paragraphs 49 and 51. Examples of sets which are everywhere discontinuous and yet whose projections on a particular line fill a segment are well-known. See Encyclopädie Math., II 2, p. 933.

[13]) [12], Theorem IX″ (p. 213).

ends of the second kind every pair of which is separated by a circle through O which meets the frontier of D in no other point. Removal of O thus renders the remainder of the set $\Gamma_F(f)$ of accessible points totally disconnected.

This example, for which $D(f)=\varkappa$, also illustrates the extremal case of Theorem 1. For, by a theorem of COLLINGWOOD[14]), the set $D_1(f)$ of points of $\varkappa$ corresponding to prime ends of the first kind is residual, i.e the complement $\mathfrak{C}D_1=\varkappa-D_1=D-D_1$ is of category I. But in our example $\Gamma_F(f, D_1)$ consists of the single point O so that it is of zero measure, while $\Gamma_F(f, D-D_1)$ is an infinite set whose projection fills the diameter of a circle through O and is thus of positive measure. Further, by the theorem of F. and M. RIESZ, $D_1(f)$ is of measure zero.

References

[1] COLLINGWOOD, E. F.: On the linear and angular clusters sets of functions meromorphic in the unit circle. Acta math. **91**, 165–185 (1954).

[2] — On sets of maximum indetermination of analytic functions. Math. Z. **67**, 377–396 (1957).

[3] — Cluster sets and prime ends. Ann. Acad. Sci. Fenn. Ser. A 1 250–256 (1958).

[4] — A theorem on prime ends. J. London Math. Soc. **31**, 344–349 (1956).

[4a] — On functions meromorphic in the unit disc and restricted on a spiral converging to the boundary. J. Indian Math. Soc. **24**, 223–229 (1961).

[5] —, and M. L. CARTWRIGHT: Boundary theorems for a function meromorphic in the unit circle. Acta math. **87**, 83–146 (1952).

[6] —, and A. J. LOHWATER: Applications of the theory of cluster sets to a class of meromorphic functions. Proc. Cambridge Phil. Soc. **53**, 93–105 (1957).

[7] NEVANLINNA, R.: Eindeutige analytische Funktionen, 2nd. edit. Berlin 1953.

[8] PRIVALOV, I. I.: Sur quelques applications de la mesure harmonique des ensembles de points à certains problèmes de la théorie des fonctions. Rec. Math. Moscou, N.s. **3**, 527–532 (Russian; French summary 533) (1938).

[9] — Randeigenschaften analytischer Funktionen. Berlin 1956.

[10] TSUJI, M.: Theory of meromorphic functions in a neighbourhood of a closed set of capacity zero. Jap. J. Math. **19**, 93–105 (1944–1948).

[11] CARATHÉODORY, C.: Über die Begrenzung einfach zusammenhängender Gebiete. Math. Ann. **73**, 323–370 (1913).

[12] KNASTER, B., and C. KURATOWSKI: Sur les ensembles connexes. Fund. Math. **2**, 206–255 (1921).

[13] SEIDEL, W.: Holomorphic functions with spiral asymptotic paths. Nagoya Math. J. **14**, 159–171 (1959).

[14]) [4], Theorem 4 (p. 349).

Girton College, Cambridge.

Lilburn Tower, Alnwick, Northumberland

(Eingegangen am 7. Dezember 1960)

WAGNER, A.
Math. Zeitschr. 76, 411–426 (1961)

On collineation groups of projective spaces. I

By
A. WAGNER

§ 1. Introduction

In this paper we shall deal mainly with the following question: what are the collineation groups of a finite projective space doubly transitive on the points of the space?

Our main results are: Theorem 2, that all such groups contain doubly transitive simple subgroups, and Theorem 4, that if the dimension of the space is less than 5, then, except for the 15-point space (*i.e.* the three-dimensional space with three points on every line), all doubly transitive groups contain the little projective group of the space (*i.e.* the unimodular projective group).

A collineation group containing the little projective group is always doubly transitive on the space. That the converse is not always true is shown by the case of the 15-point space. Whether this is the only exception remains an open question. However, it seems very likely that this is not the only exception and that other examples will occur already in spaces of dimension 5.

We may remark in passing that for the case of finite projective spaces with 3 points on every line, our question is equivalent to the question of classifying all triply transitive permutation groups with normal subgroups not more than singly transitive.

In § 2 we discuss the basic concepts involved. In § 3 we list results which will be needed later. Some of these are straightforward generalisations of known results for finite projective planes, and only brief indications are given of the proof. In § 4 we obtain some useful preliminary results concerning collineation groups containing perspectivities. In § 5 we prove our main results on doubly transitive collineation groups. The exceptional case of the 15-point space is discussed in Appendix I and is used in Appendices II and III as a source for counter-examples.

The author wishes to express his indebtedness to Professor MARSHALL HALL for drawing his attention to the question with which this paper mainly deals.

§ 2. Definitions and basic concepts

The terms *projective space* and *linear subspace, subspace* for short, will be used in the usual sense and for our purpose may most conveniently be thought of as being defined synthetically, see for example VEBLEN and YOUNG [*10*]. *Points* are subspaces of dimension 0, *lines* subspaces of dimension 1, *planes* subspaces of dimension 2, and, if the space has dimension k,

hyperplanes are subspaces of dimension $k-1$. Normally a projective space will be understood to have dimension at least 2. We shall denote a projective space of dimension k by S_k. If $k \geqq 3$ or the space is a desarguesian plane, then S_k may be co-ordinatized by a skew field. If the space is pappian, then S_k may be co-ordinatized by a field.

A *collineation* of S_k is a permutation of the subspaces preserving their dimension and incidence. If S_k may be co-ordinatized by a field $\mathfrak{F}$, then the group of all collineations of S_k is isomorphic to the quotient group of the group of non-singular semi-linear transformations of a $k+1$ dimensional vector space over $\mathfrak{F}$ by its centre.

A collineation σ of S_k fixing all points of a hyperplane H_{k-1} is called a *perspectivity*. σ will also fix all hyperplanes through some point P_0. P_0 and H_{k-1} will be called the *centre* and *axis* of σ respectively. We shall also say that σ is a (P_0, H_{k-1})-perspectivity. If $P_0 \in H_{k-1}$, σ will be called an *elation*, and if $P_0 \notin H_{k-1}$, σ will be called a *homology*.

The group of collineations generated by all perspectivites of S_k will be called the *full projective group*. If S_k may be co-ordinatized by a field $\mathfrak{F}$, then the full projective group of S_k is isomorphic to the quotient group of the group of non-singular linear transformations of a $k+1$ dimensional vector space over $\mathfrak{F}$ by its centre. This group is also called the *projective linear group*.

The group of collineations generated by all elations of S_k will be called the *little projective group*. If S_k is co-ordinatized by a field $\mathfrak{F}$, then this group is isomorphic to the quotient group of the group of unimodular linear transformations of a vector space of dimension $k+1$ over $\mathfrak{F}$ by its centre. This group is also called the *unimodular projective group*. It is also isomorphic to the subgroup of the projective linear groups consisting of those elements with determinant a $(k+1)$-th power in $\mathfrak{F}$.

For a treatment of the above topics, see for example ARTIN [2].

As is usual, subspaces of S_k of dimension s and $k-s-1$ will be considered *dual*. A one-to-one mapping of the linear subspaces of S_k onto their duals and preserving incidence will be called a *correlation*. A correlation of order 2 will be called a *polarity*. If ϑ is a polarity and P_0 a point of S_k such that $P_0 \in P_0\vartheta$, then P_0 will be called an *absolute point* of ϑ.

In this paper we shall consider only *finite projective spaces*, *i.e.* spaces with a finite number of points. If a line contains $n+1$ points, so do all lines, and n will be called the *order* of the space. If the space may be co-ordinatized by a skew field $\mathfrak{F}$, then $\mathfrak{F}$ contains n elements. Since $\mathfrak{F}$ is finite it will be a field. Hence in this case n is a power of a prime. If S_k has order n, then S_k contains $(n^{k+1}-1)/(n-1)$ points. This number we shall often denote by n_k.

A k-fold transitive permutation group will be said to be *sharply* k-fold transitive if only the identity fixes k symbols.

With the hope of facilitating a reading of this paper, we shall now make some more informal comments.

For the case of finite projective planes, the results of this paper are contained in results by OSTROM and WAGNER [8], [9]. However, for the sake of completeness, the results here are stated to include the case of planes. Since we are thus dealing essentially with spaces of dimension greater than 2, all spaces may be thought of as being co-ordinatized by a field.

The arguments of Theorems 3 and 4 are made more complicated by the need to distinguish between the full and the little projective groups. For this purpose, appeal is made to determinants in Lemma 4. Otherwise the argument makes no overt use of a co-ordinatizing field and is based only on geometrical and permutation group theoretical concepts. It is for this reason that we have used the geometrical names full and little projective group in preference to their better known algebraic equivalents.

If P_0 is any point and H_{k-1} any hyperplane of S_k and L_1 any line through P_0, then the group of all (P_0, H_{k-1})-perspectivities is transitive on the points of L_1 with P_0 and $H_{k-1} \cap L_1$ omitted. Thus S_k contains "all possible" perspectivities. The order of the full projective group may be readily computed and is

$$n_k \cdot n \cdot n_{k-1} \cdot n^2 \cdot n_{k-2} \cdot \cdots \cdot n^k \cdot (n-1)^k,$$

where n is the order of S_k. The order of the little projective group is the order of the full projective group divided by d, where $d=(k+1, n-1)$ — see Lemma 4. (See DICKSON [5], pp. 77 and 87.) When $n=2$, the collineation group, the full and little projective groups all coincide. Also there are no homologies.

The little projective group cannot be defined geometrically for a projective line. To avoid re-phrasing, we shall use the geometric name for the corresponding projective unimodular group.

To avoid tedious repetition in the proofs, all perspectivities will normally be understood to differ from the identity. Also all collineations appearing in the proofs will be understood to belong to the group stated in the enunciation of the result.

If Π is any permutation group, then $\Pi(S, T, \ldots)$ will denote the subgroup of Π consisting of those elements of Π fixing the symbols $S, T, \ldots$.

Finally, there are two simple results which are used repeatedly in the proofs.

(i) If σ is a (P_0, H_{k-1})-perspectivity and α is any collineation, then $\alpha^{-1}\sigma\alpha$ is a $(P_0\alpha, H_{k-1}\alpha)$-perspectivity.

(ii) If Π is a permutation group and $\mathfrak{S}$ a transitive class of s symbols under Π, then the order of $\Pi = s \cdot$ order of $\Pi(S)$, where S is any symbol of $\mathfrak{S}$.

§ 3. Results required later

I. Let Π be a permutation group and $\mathfrak{S}$ and $\mathfrak{T}$ two transitive classes under Π If the number of elements in $\mathfrak{S}$ is prime to the number of elements in $\mathfrak{T}$ and S is any element of $\mathfrak{S}$, then $\Pi(S)$ is transitive on $\mathfrak{T}$.

II. Let Π be a permutation group on a set $\mathfrak{S}$ and let Σ be a normal subgroup of Π with $\Sigma \neq 1$.

(a) If Π is transitive on $\mathfrak{S}$, then Σ splits $\mathfrak{S}$ into transitive classes each containing the same number of elements.

(b) If Π is doubly transitive on $\mathfrak{S}$, then Σ is transitive on $\mathfrak{S}$.

(c) If Π is triply transitive on $\mathfrak{S}$, and Σ is not sharply transitive on $\mathfrak{S}$, then Σ is doubly transitive on $\mathfrak{S}$.

The next three results are true for all finite λ-spaces, hence finite projective spaces. Proofs of these results are given in DEMBOWSKI [4].

III. Let γ be a collineation of a finite projective space. Then the number of points left fixed by γ is equal to the number of hyperplanes left fixed by γ.

IV. Let Γ be a collineation group of a finite projective space. Then the number of transitive classes of points under Γ is equal to the number of transitive classes of hyperplanes under Γ.

V. If Γ is a collineation group of a finite projective space S_k doubly transitive on the points of S_k, then Γ is also doubly transitive on the hyperplanes of S_k.

In view of Results IV and V, we shall refer to collineation groups which are transitive (doubly transitive) on the points of a space simply as *transitive (doubly transitive)* collineation groups.

VI. Let Γ be a collineation group of a finite projective space S_k. Further, for every point P_0 lying on a given hyperplane H_{k-1}, the group of all (P_0, H_{k-1})-elations contained in Γ has the same order greater than 1. Then the group of all elations with axis H_{k-1} contained in Γ is transitive on the points of S_k not lying on H_{k-1} (*i.e.* Γ contains all elations with axis H_{k-1}).

For $k=2$, *i.e.* finite projective planes not necessarily desarguesian, this result was proved by GLEASON [6]. Gleason's proof may readily be extended to higher dimensional spaces, see WAGNER [12]. Also the analogous result for λ-spaces has been proved by LÜNEBURG [7], and this contains the above as a special case.

VII. Let S_k be a finite projective space and let α and β be two homologies with centres A_0 and B_0 respectively and common axis H_{k-1}. Then the group generated by α and β contains an elation with axis H_{k-1} mapping A_0 onto B_0.

For $k=2$, again without assuming that the plane is desarguesian, this result was proved by ANDRÉ [1]. His proof may be extended to higher dimensions. The result may also be deduced immediately from the theory of Frobenius permutation groups, see WAGNER [12].

VIII. If ϑ is a polarity of a finite projective space S_k, then ϑ possesses absolute points.

For $k=2$, this result was proved by BAER [3]. For higher dimensions this result may either be proved algebraically or deduced from BAER's result by induction. We outline the latter proof.

We assume the result true for dimension $k-1$ but false for S_k. Let H_{k-1} be a hyperplane of S_k. If P_0 is any point on H_{k-1}, let $P_{k-1}=P_0\vartheta$. In view of our assumption, $P_{k-1}\neq H_{k-1}$. Now let $P_{k-2}=H_{k-1}\cap P_{k-1}$. Also, by our assumption, $P_0\notin P_{k-2}$. Further, if Q_0 is any other point of H_{k-1} distinct from P_0, $Q_{k-2}\neq P_{k-2}$ since otherwise $H_{k-1}\vartheta\in P_{k-2}$. Consequently the mapping $P_0\to P_{k-2}$ is a one-to-one mapping of the points of H_{k-1} onto the hyperplanes of H_{k-1}. It is easily seen that this is a polarity which by our assumption has no absolute points. This contradiction establishes the result.

IX. Let Π be a collineation group of a finite projective plane and let Π be transitive on non-incident point-line pairs. Then the plane is desarguesian and Π contains the little projective group of the plane.

It was shown by OSTROM [8] that Π is doubly transitive on the plane, and for Π doubly transitive the above result was proved by OSTROM and WAGNER [9].

X. If Π is a collineation group of a projective space S_k and the normaliser of Π contains the little projective group of S_k, then Π contains the little projective group of S_k.

This result is proved in ARTIN [2], p. 165, with the assumption that Π is contained in the full projective group. This assumption, however, is not necessary to the proof.

§ 4. Collineation groups containing perspectivities

LEMMA 1. *Let Π be a collineation group of S_k with the properties:*

(i) *Π is transitive on S_k.*

(ii) *Π contains a non-trivial homology.*

Then Π contains an elation.

It will be convenient to call a hyperplane P_{k-1} *an axis of the point* P_0 if there exists a (P_0, P_{k-1})-homology. Also we shall call P_0 a *centre of the hyperplane* P_{k-1}.

Since Π is transitive on S_k, every point of S_k is the centre of a homology and every hyperplane of S_k is the axis of a homology. If a hyperplane has more than one centre, the lemma follows from Result VII. Similarly, if a point has more than one axis, the lemma follows from the dual of Result VII. Thus we need only consider the case where there is a one-to-one correspondence of points and hyperplanes determined by the relationship "axis" and "centre". It is easily seen that this correspondence is a polarity with no absolute points. A proof of this assertion is given for the case of planes in WAGNER [11] (Statement I of Theorem 2), and is directly applicable to higher dimensions. In view of Result IV, such a polarity is impossible, and the contradiction establishes the lemma.

LEMMA 2. *Let Π be a doubly transitive collineation group of a finite projective space S_k. Then if H_{k-1} and K_{k-1} are any two hyperplanes of S_k and P_0 is any point of S_k, with $P_0\in K_{k-1}$ and $P_0\notin H_{k-1}$, and if H'_{k-1}, K'_{k-1}, P'_0*

are defined similarly, there exists a collineation $\pi \in \Pi$ *such that* $H_{k-1}\pi = H'_{k-1}$, $K_{k-1}\pi = K'_{k-1}$ *and* $P_0\pi = P'_0$.

Let L_{k-1} be any hyperplane of S_k. Then $\Pi(L_{k-1})$ splits the hyperplanes of S_k into two transitive classes, namely L_{k-1} and all the other hyperplanes. Hence by Result IV $\Pi(L_{k-1})$ splits the points of S_k into two transitive classes, namely the points of L_{k-1} and the points of S_k not on L_{k-1}. Now since $\Pi(L_{k-1})$ is transitive on the points of L_{k-1}, it is transitive on the hyperplanes of L_{k-1}. The number of hyperplanes of L_{k-1} is n_{k-1} and the number of points of S_k not on L_{k-1} is n^k. Since $(n^k, n_{k-1}) = 1$, it follows from Result I that the subgroup of $\Pi(L_{k-1})$ fixing a given hyperplane of L_{k-1} is still transitive on the points of S_k not on L_{k-1}.

To complete the proof of the lemma, we first observe that there exists an element $\alpha \in \Pi$ with $H_{k-1}\alpha = H'_{k-1}$. Clearly $P_0\alpha \notin H'_{k-1}$. Also let $K_{k-1}\alpha \cap H'_{k-1} = K_{k-2}$ and $K'_{k-1}\alpha \cap H'_{k-1} = K'_{k-2}$. Then from the above it follows that there exists an element $\beta \in \Pi(H'_{k-1})$ which maps K_{k-2} onto K'_{k-2} and $P_0\alpha$ onto P'_0. Then $\pi = \alpha\beta$ is the required collineation.

We may restate Lemma 2 in a weaker form as follows:

COROLLARY. *Let* Π *be a doubly transitive collineation group of* S_k *and let* H_{k-1} *be any hyperplane of* Π. *Then* $\Pi(H_{k-1})$ *is transitive on the points of* S_k *not on* H_{k-1} *and also* $\Pi(H_{k-1})$ *is transitive on non-incident point-hyperplane pairs of* H_{k-1}.

THEOREM 1. *Let* Π *be a collineation group of a finite projective space* S_k. *Then if* Π *contains a non-trivial elation (homology) and if* Π *is transitive on incident (non-incident) point-hyperplane pairs of* S_k, Π *contains the little projective group of* S_k.

(i) We first consider the case where Π contains an elation and is transitive on incident point-hyperplane pairs. For all points P_0 and hyperplanes H_{k-1} in S_k, the groups of (P_0, H_{k-1})-elations in Π are all conjugate in Π, hence have the same order greater than 1. If we first fix H_{k-1}, we obtain from Result VI all elations with axis H_{k-1}, hence by taking all hyperplanes of S_k it follows that Π contains all elations of S_k.

(ii) We now consider the case where Π contains a homology and is transitive on non-incident point-hyperplane pairs in S_k. In this case an argument analogous to the above with an appeal to Result VII proves the theorem.

From Lemma 1 and this theorem we obtain immediately the following corollary.

COROLLARY 1. *If* Π *is a collineation group of a finite projective space* S_k *transitive on incident point-hyperplane pairs and if* Π *contains a homology, then* Π *contains the little projective group of* S_k.

Also, by combining Theorem 1 with Lemma 2, we obtain immediately the following corollary.

COROLLARY 2. *If* Π *is a doubly transitive collineation group of a finite projective space* S_k *and if* Π *contains a non-trivial perspectivity, then* Π *contains the little projective group of* S_k.

§ 5. Doubly transitive collineation groups

The following lemma is important for a more detailed knowledge of doubly transitive collineation groups.

LEMMA 3. *Let Π be a doubly transitive collineation group of a finite projective space S_k and let P_2 be any plane of S_k. Then $\Pi(P_2)$, as a collineation group of P_2, contains the little projective group of P_2.*

When $k=2$ this lemma follows from Result IX. We shall henceforth assume $k \geqq 3$.

Let M_{k-3} be any subspace of dimension $k-3$ in S_k, let L_{k-2} be any subspace of dimension $k-2$ containing M_{k-3} and let H_{k-1} be any hyperplane containing L_{k-2}. Now consider the group $\Pi(H_{k-1}, L_{k-2})$. In view of the corollary of Lemma 2, this group is transitive on the points of H_{k-1} not on L_{k-2}. Let this group split the points of L_{k-2} into s transitive classes. Thus $\Pi(H_{k-1}, L_{k-2})$ splits the points of H_{k-1} into $1+s$ transitive classes, hence also the hyperplanes of H_{k-1} into $1+s$ transitive classes. Then the hyperplanes of H_{k-1} distinct from L_{k-2} split into s transitive classes. Now regarding $\Pi(H_{k-1}, L_{k-2})$ as a collineation group of L_{k-2}, it follows that this group splits the hyperplanes of L_{k-2} into s transitive classes. Thus all the hyperplanes of H_{k-1} containing M_{k-3} and distinct from L_{k-2} belong to one transitive class. Hence $\Pi(H_{k-1}, L_{k-2}, M_{k-3})$ is transitive on the hyperplanes of H_{k-1} distinct from L_{k-2} and containing M_{k-3}.

Since double transitivity is a dual property, we may dualize the last result to obtain:

If P_2 is any plane of S_k and Q_1 any line of P_2 and R_0 any point of Q_1, then $\Pi(P_2, Q_1, R_0)$ is transitive on the lines through R_0 distinct from Q_1 in P_2. The following steps are easily verified:

(i) If R_0 is any point of P_2, then $\Pi(P_2, R_0)$ is doubly transitive on the lines of P_2 through R_0.

(ii) $\Pi(P_2)$ is transitive on the lines of P_2.

(iii) $\Pi(P_2)$ is transitive on the points of P_2.

(iv) $\Pi(P_2)$ is doubly transitive on the lines of P_2.

The lemma now follows by using Result IX.

COROLLARY. *If Q_1 is any line of S_k, then $\Pi(Q_1)$ is triply transitive on the points of Q_1.*

THEOREM 2. *Let Π be a doubly transitive collineation group of a finite projective space S_k. Then any normal subgroup of Π, not the identity, is also doubly transitive on S_k.*

When $k=2$ it follows from Result IX that Π contains the little projective group of the plane, and the theorem now follows from Result X.

Henceforth we shall assume that $k \geqq 3$. Also let n be the order of S_k and let Σ be a normal subgroup of Π, with $\Sigma \neq 1$. The proof will be carried out in a number of steps. We shall make repeated use of result II.

Step I. n_k is not a power of 2.

When n is even or when n is odd and k is even, n_k is odd. Hence it only remains to consider the case n is odd and k is odd. Now $n_3=n^3+n^2+n+1=(n^2+1)(n+1)=\{(n+1)^2-2n\}\{n+1\}$ and it is easily seen that this cannot be a power of 2. Now let $l=2m-1$ be the least integer greater than 3 for which n_k is a power of 2. $n_l=(n^{2m}-1)/(n-1)=(n^m+1)(n^m-1)/(n-1)=(n^m+1)n_{m-1}$, hence n_{m-1} is a power of 2. Since n_2 is odd, $l\neq 5$. We have shown that n_3 is not a power of 2, hence $l\neq 7$. If $l\geqq 9$, then $m-1>3$, contradicting our assumption that l is the least integer greater than 3 for which n_l is a power of 2.

Step II. 2_k, $k>1$, is not a power of 3.

$2_k=2^{k+1}-1$, hence if 2_k is a power of 3, then $2^{k+1}\equiv 1\ (3)$ and k is odd. Let $l=2m-1$ be the least odd integer greater than 1 for which 2_k is a power of 3. Now $2_l=2^{2m}-1=(2^m+1)(2^m-1)=(2^m+1)2_{m-1}$, hence 2_{m-1} is a power of 3. Now $2_3=15$, hence $l>3$. If $l>3$, then $m-1>1$, contradicting our assumption that l is the least integer greater than 1 for which 2_l is a power of 3.

Step III. Σ is not sharply transitive on S_k.

Since Π is doubly transitive on S_k, it follows from result II that Σ is transitive on the points of S_k. We wish to show that if S_0 is any point of S_k, then $\Sigma(S_0)\neq 1$.

Since n_k is not a power of 2, the order of Σ is not a power of 2, hence Σ contains an element σ of odd order. Let $T_0=S_0\sigma$ and $U_0=T_0\sigma$. We shall deal separately with the case S_0, T_0, U_0 non-collinear and the case S_0, T_0, U_0 collinear.

We must first make an observation regarding the case $n=2$. Since n_k is odd, it follows from Step II that n_k has a prime factor greater than 3, say s. Now let σ be an element of Σ of order s. If S_0, T_0, U_0 are defined as before, then S_0, T_0, U_0 cannot be collinear, since otherwise σ fixes the line S_0, T_0, U_0 and the order of σ would be 2 or 3.

(i) *S_0, T_0, U_0 non-collinear.*

Let P_2 be the plane containing these points. Then by Lemma 3 there exists an element $\pi\in\Pi(P_2)$ which fixes S_0 and T_0 but does not fix U_0. Now consider the element $\varrho=\pi\sigma\pi^{-1}\sigma^{-1}$. Clearly $\varrho\in\Sigma$, also $S_0\varrho=S_0$, hence $\varrho\in\Sigma(S_0)$. Also $T_0\varrho\neq T_0$, hence $\Sigma(S_0)\neq 1$.

(ii) *S_0, T_0, U_0 collinear.*

Let Q_1 be the line containing these points. We may assume $n>2$. It follows from Lemma 3 that $\Pi(Q_1)$ contains an element π such that π fixes S_0 and T_0 but does not fix U_0. Now, as in case (i), consideration of $\pi\sigma\pi^{-1}\sigma^{-1}$ establishes the result.

Consideration of the order of Σ shows that if Σ is not sharply transitive on the points of S_k, Σ is not sharply transitive on the hyperplanes of S_k.

Step IV. For all $S_{k-2} \in S_k$, $\Sigma(S_{k-2})$ is doubly transitive on the hyperplanes containing S_{k-2}.

Let H_{k-1} be any hyperplane of S_k. $\Sigma(H_{k-1})$ is a normal subgroup of $\Pi(H_{k-1})$. Since, by Lemma 3, $\Pi(H_{k-1})$ is transitive on the points (hyperplanes) of H_{k-1}, it follows that $\Sigma(H_{k-1})$ will split the points (hyperplanes) of H_{k-1} into, say, r transitive classes each with the same number of elements, say u where $u \mid n_{k-1}$. Similarly, since $\Pi(H_{k-1})$ is transitive on the points not on H_{k-1}, $\Sigma(H_{k-1})$ will split the points not on H_{k-1} into, say, s transitive classes each with the same number of elements, say v where $v \mid n^k$.

Since $\Sigma(H_{k-1}) \neq 1$, it follows that $v > 1$. Clearly $(v, u) = 1$. Let U_{k-2} be any hyperplane of H_{k-1}. Then $\Sigma(H_{k-1}, U_{k-2})$ will by result I still split the points not on H_{k-1} into s transitive classes, each with v elements.

Now let P_0 and Q_0 be any two points not on H_{k-1} which lie in the same transitive class under $\Sigma(H_{k-1})$. Let L_{k-1} be a hyperplane such that $P_0 \in L_{k-1}$ and $Q_0 \notin L_{k-1}$ and let $R_{k-2} = L_{k-1} \cap H_{k-1}$. Then $\Sigma(H_{k-1}, R_{k-2})$ will by our previous observation still map P_0 onto Q_0, hence will not leave L_{k-1} fixed. Therefore $\Sigma(R_{k-2})$ considered as a permutation group on the hyperplanes containing R_{k-2} is not the identity. The dual of the corollary of Lemma 3 states $\Pi(R_{k-2})$ is triply transitive on the hyperplanes containing R_{k-2}. Hence $\Sigma(R_{k-2})$ is transitive on the hyperplanes containing R_{k-2}, and the above argument has shown that $\Sigma(R_{k-2})$ is not sharply transitive, hence $\Sigma(R_{k-2})$ is doubly transitive on the hyperplanes containing R_{k-2}. Since Π is transitive on the S_{k-2} of S_k and Σ is normal in Π, Step IV follows by taking conjugates of $\Sigma(R_{k-2})$ in Π.

Step V. If H_{k-1} is any hyperplane of S_k, $\Sigma(H_{k-1})$ is transitive on the points not lying on H_{k-1}.

Using the notation of Step IV, $\Sigma(H_{k-1})$ splits the points of S_k into $r+s$ transitive classes. Now if R_{k-2} is any hyperplane of H_{k-1} we have shown that $\Sigma(H_{k-1})$ is transitive on all hyperplanes of S_k containing R_{k-2} and distinct from H_{k-1}. Hence $\Sigma(H_{k-1})$ splits the hyperplanes distinct from H_{k-1} into r transitive classes, hence splits the hyperplanes of S_k into $r+1$ transitive classes. Since by Result IV $r+s = r+1$, it follows that $s = 1$, *i.e.* that $\Sigma(H_{k-1})$ is transitive on the points not lying on H_{k-1}.

Step VI. If H_{k-1} is any hyperplane of S_k and R_{k-2} is any hyperplane of H_{k-1}, then $\Sigma(H_{k-1}, R_{k-2})$ is transitive on the points of H_{k-1} not lying on R_{k-2}.

Let L_{k-1} be any hyperplane of S_k such that $R_{k-2} = L_{k-1} \cap H_{k-1}$. Then $\Sigma(L_{k-1})$ is transitive on the points not lying on L_{k-1}, hence, as in Step IV, it follows that $\Sigma(L_{k-1}, R_{k-2})$ is also transitive on the points not lying on L_{k-1}. Now if P_0 and Q_0 are any two points of H_{k-1} not lying in L_{k-1}, there exists an element $\sigma \in \Sigma(L_{k-1}, R_{k-2})$ such that $P_0\sigma = Q_0$. Clearly σ fixes H_{k-1}, hence $\sigma \in \Sigma(L_{k-1}, H_{k-1})$. Therefore $\Sigma(L_{k-1}, H_{k-1})$ has the desired property, hence also the larger group $\Sigma(H_{k-1}, R_{k-2})$.

Step VII. Σ is transitive on all S_{k-2} in S_k.

Let H_{k-1} be any hyperplane of S_k and let P_0 and Q_0 be any two distinct points of H_{k-1}. Further, let R_{k-2} be a hyperplane of H_{k-1} such that $P_0, Q_0 \notin R_{k-2}$. By Step VI there exists an element of $\Sigma(H_{k-1}, R_{k-2})$ mapping P_0 onto Q_0, hence $\Sigma(H_{k-1})$ is transitive on the points of H_{k-1}, hence also on the hyperplanes of H_{k-1}. Since Σ is transitive on the hyperplanes of S_k, it immediately follows that Σ is transitive on all $S_{k-2} \in S_k$.

Theorem 2 is now an immediate consequence of combining Steps IV and VII.

Any collineation group of a projective space of dimension greater than 2 may be represented by semilinear transformations. The group of automorphisms of a finite field is well-known to be cyclic, and a little computation shows that the commutator subgroup of any group of semilinear transformations over a finite field will consist only of linear transformations, in fact only of unimodular linear transformations. Since every doubly transitive collineation group is certainly non-abelian, its commutator subgroup is not the identity, hence by Theorem 2 is also doubly transitive. Consequently it is henceforth generally sufficient to consider the case of doubly transitive collineation groups which are subgroups of the little projective group.

The following lemma will be required in the proof of the next theorem.

LEMMA 4. *Let Σ be the little projective group of a finite projective space S_k of order n. Then Σ contains non-trivial homologies if and only if $n-1 \mid k+1$. Also Σ is the full projective group of S_k if and only if $(n-1, k+1)=1$.*

Let $\mathfrak{F}$ be the field with n elements and let ω be a generator of the multiplicative group of $\mathfrak{F}$. ω has order $n-1$. Now, by a suitable choice of basis in S_k, any homology of S_k may be represented by a $(k+1)\times(k+1)$ diagonal matrix $\boldsymbol{H}$ with diagonal entries $\omega^i, 1, 1, \ldots, 1$ for some integer i. Now $\boldsymbol{H} \in \Sigma$ if and only if there exists a ω^s such that $\omega^s \boldsymbol{H}$ is unimodular, *i.e.* if and only if $\omega^{s(k+1)+i}=1$. Hence Σ contains homologies if, and only if, $s(k+1)+i \equiv 0 (n-1)$ has solutions other than $i \equiv 0 (n-1)$. The lemma is now a restatement of this equation.

THEOREM 3. *Let Π be a collineation group of a finite projective space S_k, $k \geqq 3$, of order n with the following properties:*

(i) *Π is transitive on S_k.*

(ii) *If S_{k-1} is any hyperplane of S_k, then $\Pi(S_{k-1})$, as a collineation group of S_{k-1}, contains the little projective group of S_{k-1}.*

Then, unless $k=3$ and $n=2$, Π contains the little projective group of S_k.

It is an immediate consequence of (i) and (ii) that Π is doubly transitive on S_k. If $\Pi(S_{k-1})$ is not faithful on S_{k-1}, it follows from Corollary 2 of Theorem 1 that Π contains the little projective group of S_k. Consequently we shall henceforth assume that $\Pi(S_{k-1})$ is faithful on S_{k-1}.

If Π' is the commutator subgroup of Π, it follows from Theorem 2 that Π' is doubly transitive on S_k, hence has property (i). Since $\Pi'(S_{k-1}) \neq 1$, it follows from Result X that $\Pi'(S_{k-1})$ has property (ii).

Thus without loss of generality we may assume that Π is a subgroup of the little projective group of S_k. This fact will be used only in Step VI.

We shall break up the proof into a number of steps.

Step I. If P_0 is any point of S_k not on S_{k-1} and if $\Pi(S_{k-1}, P_0)$ contains the little projective group of S_{k-1}, then Π contains the little projective group of S_k.

Let $\Pi^*(S_{k-1})$ be the subgroup of $\Pi(S_{k-1})$ consisting of those elements of $\Pi(S_{k-1})$ which as collineations of S_{k-1} belong to the little projective group of S_{k-1}. $\Pi^*(S_{k-1})$ is clearly normal in $\Pi(S_{k-1})$. Since we are assuming that $\Pi(S_{k-1})$ is faithful on S_{k-1}, it follows that $\Pi^*(S_{k-1}) = \Pi^*(S_{k-1}, P_0)$.

Now let Q_0 be any other point of S_k not on S_{k-1}. By the corollary of Lemma 2 there exists an element $\pi \in \Pi(S_{k-1})$ such that $P_0\pi = Q_0$. Since $\pi^{-1}\Pi^*(S_{k-1}, P_0)\pi = \Pi^*(S_{k-1}, Q_0)$ and $\pi^{-1}\Pi^*(S_{k-1})\pi = \Pi^*(S_{k-1})$, it follows that $\Pi^*(S_{k-1})$ fixes both P_0 and Q_0, hence all points of S_k not on S_{k-1}, which is impossible.

Step II. Let P_0 be any point of S_k not on S_{k-1}. Then $\Pi(S_{k-1}, P_0)$ is transitive on incident point-hyperplane pairs of S_{k-1}.

Using the corollary of Lemma 2, it follows that P_0 lies in a transitive class of n^k points under $\Pi(S_{k-1})$. Also $\Pi(S_{k-1})$ is transitive on the $n_{k-1}n_{k-2}$ incident line-hyperplane pairs of S_{k-1}. Since $(n^k, n_{k-1}n_{k-2}) = 1$, Step II now follows from Result I.

Step III. If P_0 is any point of S_k not on S_{k-1} and $\Pi(S_{k-1}, P_0)$ contains a perspectivity of S_{k-1}, then Π contains the little projective group of S_k.

Using Step II and Theorem 1 and Corollary 1, it follows that $\Pi(S_{k-1}, P_0)$ contains the little projective group of S_{k-1}. Step III now follows from Step I.

In the next step we include the line among the projective spaces.

Step IV. If Σ is the little projective group of a projective space S_k of order n, and $n \leqq k+3$, then, unless $k=1$ and $n=2$ or 3, Σ does not contain a subgroup of index n.

We shall first deal with the case $k=1$. $n=2$ or 3 have been excluded. When $n=4$, Σ is the simple group of order 60 and this is well-known not to have a subgroup of order 15. We shall now consider the case $k \geqq 2$. In this case Σ is always simple. The order of Σ is $n_k n n_{k-1} n^2 n_{k-2} \cdots n^k (n-1)^k/d$ where d is a divisor of $n-1$. Since $n = p^t$ where p is a prime and $(n_i, n) = 1$, for all i, the highest power of p dividing the order of Σ is $n^{k(k+1)/2}$. Also a little computation shows that the highest power of p dividing $n!$ is $p^{p^t-1/p-1}$.

Now if Σ contains a subgroup of index n, Σ can be faithfully represented on n symbols and Σ must be a subgroup of the symmetric group on n symbols. Hence $n^{k(k+1)/2} \mid p^{p^t-1/p-1}$ *i.e.* $tk(k+1)/2 \leqq (p^t-1)/(p-1)$. When $k=2$ this inequality is clearly false. When $k \geqq 3$ $k(k+1)/2 \geqq k+3 \geqq n \geqq p^t - 1/p - 1$ and the inequality is also false.

Step V. If $\Pi(S_{k-1})$ contains a homology of S_{k-1}, then Π contains the little projective group of S_k.

Let σ be the homology of S_{k-1}. If σ has order s as a collineation of S_{k-1}, then clearly $s \mid n-1$. Since $\Pi(S_{k-1})$ is faithful on S_{k-1}, σ has order s. There are n^k points not on S_{k-1}, and $(s, n^k)=1$, hence σ fixes at least one point not on S_{k-1}. Use of Step III now completes the proof.

Step VI. If $\Pi(S_{k-1})$ does not contain a homology of S_{k-1}, then Π contains the little projective group of S_k.

By the argument at the beginning of this theorem, we may take $\Pi(S_{k-1})$ to consist only of projective collineations.

Since $\Pi(S_{k-1})$ does not contain any homologies, the little projective group of S_{k-1}, being contained in $\Pi(S_{k-1})$, does not contain any homologies. Hence, using Lemma 4, it follows that $n-1 \mid (k-1)+1$ *i.e.* $n-1 \mid k$.

Since $n-1 \mid k$, $(n-1, k-1)=1$, hence by Lemma 4 the little projective group and the full projective group coincide for S_{k-2}. Now let S_{k-2} be a hyperplane of S_k. Since $\Pi(S_{k-1}, S_{k-2})$ contains the little projective group of S_{k-2}, $\Pi(S_{k-1}, S_{k-2})$ contains the full projective group of S_{k-2}. Let the order of the full projective group of S_{k-2} be q. The subgroup of $\Pi(S_{k-1}, S_{k-2})$ fixing S_{k-2} pointwise is the group of perspectivities of S_{k-1} with axis S_{k-2}. There are no homologies; however the elations are transitive on the n^{k-1} points of S_{k-1} not in S_{k-2}, hence the order of $\Pi(S_{k-1}, S_{k-2})=n^{k-1}q$. Now by Lemma 2 $\Pi(S_{k-1}, S_{k-2})$ is transitive on the n^k points of S_k not on S_{k-1}. Hence if P_0 is a point not on S_{k-1}, the order of $\Pi(S_{k-1}, S_{k-2}, P_0)$ is q/n. If this group is not faithful on S_{k-2}, then the group contains an elation of S_{k-1} with axis S_{k-2} and the result follows from Step III.

We now assume that $\Pi(S_{k-1}, S_{k-2}, P_0)$ is faithful on S_{k-2}. This implies that the full projective group — in this case the same as the little projective group — of S_{k-2} has a subgroup of index n. Now we have shown that $n-1 \mid k$, therefore $n-1 \leqq k$, therefore $n \leqq (k-2)+3$. We are now able to apply Step IV to reach a contradiction. Hence $\Pi(S_{k-1}, S_{k-2}, P_0)$ is not faithful on S_{k-2} and Step VI has been proved.

Step V and Step VI exhaust the possibilities for $\Pi(S_{k-1})$ and Theorem 3 has been established.

THEOREM 4. *Let Π be a doubly transitive collineation group of a finite projective space S_k of order n, with $k \leqq 4$. Then, except when $k=3$ and $n=2$, Π contains the little projective group of S_k. When $k=3$ and $n=2$, either*

(i) *Π is isomorphic to the alternating group on 7 letters or*

(ii) *Π is the group of all collineations.*

PROOF. (1) $k=2$. In this case the theorem follows from Result IX. (2) $k=3$. Let P_2 be any plane of S_3. Then, by Lemma 3, $\Pi(P_2)$ contains the little projective group of P_2. We must now deal separately with the cases $n>2$ and $n=2$.

(a) $n>2$. Π satisfies the conditions of Theorem 3, hence Π contains the little projective group of S_3.

(b) $n=2$. If $\Pi(P_2)$ is not faithful on P_2, it follows from Corollary 2 of Theorem 1 that Π contains the little projective group of the space. In

this case the little projective group is the same as the group of all collineations (there are no automorphisms of the field and every linear transformation is unimodular).

If $\Pi(P_2)$ is faithful on P_2, then the order of $\Pi(P_2)$ is 168, the order of the collineation group of the 7-point plane. Since Π is transitive on the 15 planes of the space, it follows that the order of Π is $15.168=\frac{1}{2}7!$. Theorem 2 shows that this group is simple and it is known that up to isomorphism there exists only one group with this order. This group may be represented by the alternating group on seven letters, A_7. That a collineation group with this order does in fact exist will be shown in the Appendix.

(3) $k=4$. Let P_2 be any plane and Q_3 any 3-space of S_4, with $P_2 \in Q_3$. $\Pi(P_2)$ is clearly transitive on the n_1 3-spaces containing P_2, and by Lemma 3 $\Pi(P_2)$ is transitive on the n_2 points of P_2. Since $(n_2, n_1)=1$, it follows from Result I that $\Pi(P_2, Q_3)$ is transitive on the points of P_2. Also, by the corollary of Lemma 2, $\Pi(P_2, Q_3)$ is transitive on the points of Q_3 not contained in P_2. Thus $\Pi(P_2, Q_3)$ splits the points of Q_3 into two transitive classes, hence splits the planes of Q_3 into two transitive classes, namely P_2 and the remaining planes of Q_3. Now $\Pi(Q_3)$ is transitive on the planes of Q_3, therefore $\Pi(Q_3)$ is doubly transitive on Q_3. We must now deal separately with the cases $n>2$, $n=2$.

(a) $n>2$. It follows from 2(a) that $\Pi(Q_3)$ contains the little projective group of Q_3. Thus Π satisfies the conditions of Theorem 3 and the result follows.

(b) $n=2$. It follows from 2(b) that $\Pi(Q_3)$, as a collineation group of Q_3, is either the group of all collineations of Q_3 or $\Pi(Q_3) \cong A_7$.

If the former alternative is the case, appeal to Theorem 3 again completes the proof.

Now let us suppose that $\Pi(Q_3)$ as a collineation group of Q_3, is isomorphic to A_7. If $\Pi(Q_3)$ is not faithful on Q_3, it follows as in 2(b) above that Π contains the little projective group of S_4. If $\Pi(Q_3)$ is faithful on Q_3, then the order of Π is $31.\frac{1}{2}.7!$ However, it follows from Lemma 2 that the order of a doubly transitive collineation group is divisible by n^k, *i.e.* in this case by 2^4, and since $2^4 \nmid 31.\frac{1}{2}.7!$, it follows that $\Pi(Q_3)$ cannot be faithful on Q_3.

This completes the proof of Theorem 4.

Appendix 1

In view of Theorem 4, it remains to show that the 15-point space, *i.e.* the three dimensional projective space of order 2, does possess a doubly transitive collineation group isomorphic to A_7[1]. This we shall do in two ways. The first (Part A) uses the well-known isomorphism, due to Jordan (see, for example, DICKSON [5], p. 290), between the full projective group of

[1] That such a collineation group exists has previously been noticed, though not published, by Prof. N. ITÔ. I am indepted to Dr. B. HUPPERT for pointing this out to me.

this space and A_8. The second (Part B) establishes the result geometrically, and incidentally gives a new proof of Jordan's isomorphism.

A. The full projective group Σ of the 15-point space is isomorphic to A_8, hence has a subgroup Π isomorphic to A_7. Let $\mathfrak{C}$ be an incident point-plane pair. Since Σ is transitive on incident point-plane pairs and there are 15.7 such pairs, the order of $\Sigma(\mathfrak{C})$ is $8.6.4$. Let $\mathfrak{C}$ lie in a transitive class of r elements under Π. Then the order of $\Pi(\mathfrak{C})$ is $\frac{1}{2}.7!/r$ and since $\Pi(\mathfrak{C})$ is a subgroup of $\Sigma(\mathfrak{C})$, it follows that $\frac{1}{2}.7!/r \mid 8.6.4$. This implies that $r=15.7$, *i.e.* that Π is transitive on incident point-plane pairs. Now if P_2 is any plane of the space, then, if $\Pi(P_2)$ is not faithful on P_2, it follows from Theorem 1 that Π is the full projective group of the space, which is clearly impossible. Hence $\Pi(P_2)$ is faithful on P_2, and since the order of $\Pi(P_2)=7.6.4$, $\Pi(P_2)$ is the full projective group of P_2, hence certainly doubly transitive on P_2. Consequently Π is doubly transitive on the space.

B. We may also show directly, by a method for which I am indebted to Professor A. M. GLEASON, that the 15-point space has a doubly transitive collineation group isomorphic to A_7.

Let P_1 be a 7-point plane. We may represent this plane by

$$\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 3 & 4 & 5 & 6 & 7 & 1 \\ 4 & 5 & 6 & 7 & 1 & 2 & 3 \end{array}$$

where the symbols $1-7$ are the points and the columns the lines. The full projective group Π of collineations of this plane being simple must consist only of even permutations on the symbols $1-7$. The index of Π in A_7 is 15, hence, taking all even permutations of the symbols $1-7$, we obtain 15 planes, $P_1, \ldots, P_{15}$, of which none can be changed into another by a collineation of that plane. Clearly A_7 is a transitive permutation group on these 15 planes.

The 15 planes $P_1, \ldots, P_{15}$ will now be called "points". From the symbols $1-7$, 35 unordered triples may be formed; these will be called "lines". If the 3 symbols of a "line" are collinear in a "point", then the "lines" and "points" will be called incident. Clearly A_7 is a transitive permutation group on the "lines" and A_7 preserves incidence. We now wish to show that the system defined above forms a 15-point projective space.

We first observe that the subgroup of A_7 fixing a "point" is doubly transitive on the "lines" incident with it. Thus, without loss of generality, two concurrent "lines" may be represented by (1, 2, 4) and (1, 5, 6), which are incident with P_1. Any other "point" incident with both these "lines" would be a plane where (1, 2, 4) and (1, 5, 6) are each collinear. However, there are only two such planes, namely P_1 and a plane resulting from an odd permutation of P_1. Thus two "lines" meet in at most one "point".

A "point" is clearly incident with exactly 7 "lines". The "line" (1, 2, 4) is certainly incident with P_1 and the two other "points" resulting from P_1 by cyclic permutation of (3, 6, 7). Thus every "line" is incident with at

least three points. By considering the 7 lines incident with P_1 and using the conclusion of the last paragraph, it is seen that every "line" is incident with exactly three "points" and that every two distinct "points" are incident with a unique "line".

Since every two lines of a plane have a symbol in common, it follows that two "lines" are concurrent if and only if they have exactly one symbol in common. As before, two concurrent "lines" may be taken to be (1, 2, 4) and (1, 5, 6). Now the "lines" which meet both the "lines" in "points" distinct from P_1 are (2, 5, 7), (2, 6, 3), (4, 5, 3) and (4, 6, 7). Any two of these have one symbol in common, hence are concurrent. Thus we have shown that the above system forms a 15-point projective space, and that A_7 is a collineation group of this space. The argument of Part A may now be used to show that A_7 is doubly transitive on the space.

A_7 has index 8 in the full projective group, hence the latter may be represented as a permutation group on 8 symbols, and in view of the order of the group and the fact that it is simple, the representation is the alternating group on 8 symbols. This establishes the isomorphism of Jordan.

Appendix II

Theorem 4 is based on an inductive process. If Π is a doubly transitive collineation group of a projective space S_k and S_{k-1} is a hyperplane of S_k, then $\Pi(S_{k-1})$ is, by Lemma 2, transitive on non-incident point-hyperplane pairs of S_{k-1}. For $k=3$ Result IX shows that the latter conclusion by itself implies that $\Pi(S_{k-1})$ is doubly transitive on S_{k-1}. For $k=4$ a more complicated argument was needed to establish this result, since, as the following example shows, it is not true in general that transitivity on non-incident point-hyperplane pairs implies double transitivity.

By Appendix I the 15-point space has a collineation group Π isomorphic to A_7. Consequently Π has a subgroup Θ isomorphic to A_6. Let $\mathfrak{D}$ be a non-incident point-plane pair of the space. Π is transitive on the 15.8 pairs like $\mathfrak{D}$, hence the order of $\Pi(\mathfrak{D})$ is 21. Since the order of $\Theta(\mathfrak{D})$ must divide the order of $\Pi(\mathfrak{D})$, it is easily seen that Θ is transitive on non-incident point-plane pairs. However, Θ is not doubly transitive on the space since, by Theorem 4, Π is the group of least order with this property.

Appendix III

For projective planes a much stronger form of Theorem 1 may be proved, namely that a collineation group transitive and possessing a perspectivity must contain the little projective group, see WAGNER [11]. The following example shows that this stronger result is not true for spaces of higher dimension.

Again let Σ be the full projective group of the 15-point space. Σ is isomorphic to A_8 and let Π be a subgroup of Σ isomorphic to A_7. By Theorem 1, Π does not contain any perspectivities. In this space every perspectivity is

an elation and has order 2. Thus the perspectivities are represented by permutations on 8 symbols which fix no letter, hence by 4 interchanges. A count on the number of perspectivities and the number of permutations in A_8 consisting of 4 interchanges shows that every permutation of this kind represents a perspectivity.

Now let $A_{2,6}$ be a maximal subgroup of A_8 which splits the 8 symbols into two transitive classes, one containing 2 symbols, the other 6. The order of $A_{2,6}$ is clearly $2 \cdot 6 \cdot 5 \cdot 4 \cdot 3$. Let Ω be the subgroup of Σ isomorphic to $A_{2,6}$. By our earlier observation Ω contains perspectivities of the space; in fact a count shows that there are 15. Also simple computation analogous to that in Appendix I, Part A and Appendix II shows that Ω is transitive on the space.

References

[1] ANDRÉ, J.: Über Perspektivitäten in endlichen projektiven Ebenen. Arch. Math. **6**, 29–32 (1954).

[2] ARTIN, E.: Geometric Algebra. New York 1957.

[3] BAER, R.: Polarities in finite projective planes. Bull. Amer. Math. Soc. **52**, 77–93 (1946).

[4] DEMBOWSKI, H. P.: Verallgemeinerungen von Transitivitätsklassen endlicher projektiver Ebenen. Math. Z. **69**, 59–89 (1958).

[5] DICKSON, L. E.: Linear Groups. Dover, New York 1958.

[6] GLEASON, A. M.: Finite Fano planes. Amer. J. Math. **78**, 797–807 (1956).

[7] LÜNEBURG, H.: Zentrale Automorphismen von λ-Räumen. Arch. Math. **12**, 134–145 (1961).

[8] OSTROM, T. G.: Dual transitivity in finite projective planes. Proc. Amer. Math. Soc. **9**, 55–56 (1958).

[9] —, and A. WAGNER: On projective and affine planes with transitive collineation groups. Math. Z. **71**, 186–199 (1959).

[10] VEBLEN, O., and J. W. YOUNG: Projective Geometry I. Boston 1916.

[11] WAGNER, A.: On perspectivities of finite projective planes. Math. Z. **71**, 113–123 (1959).

[12] — To appear in Proceedings of Colloquium on Foundations of Geometry. Utrecht 1959.

Imperial College, London

(Eingegangen am 29. November 1960)

HAKEN, W.
Math. Zeitschr. 76, 427–467 (1961)

Ein Verfahren zur Aufspaltung einer 3-Mannigfaltigkeit in irreduzible 3-Mannigfaltigkeiten

Von

WOLFGANG HAKEN

Einleitung

Die vorliegende Arbeit löst die Aufgabe, eine gegebene (abgeschlossene, berandete oder unberandete) 3-Mannigfaltigkeit M^3 [1]) durch 2-Sphären in *irreduzible* 3-Mannigfaltigkeiten aufzuspalten (d.h. ein Verfahren anzugeben, nach dem diese Aufspaltung stets in einer beschränkten Anzahl von Schritten erreicht werden kann). Als *irreduzibel* bezeichnen wir nach H. KNESER [6] eine 3-Mannigfaltigkeit M^3 dann, wenn jede in $\mathring{M}^3$ [2]) liegende 2-Sphäre ein in M^3 liegendes 3-Element berandet [3]). (Die 3-Sphäre ist in diesem Sinne irreduzibel [1].) Das Homöomorphieproblem der 3-Mannigfaltigkeiten ist damit zurückgeführt auf das Homöomorphieproblem der irreduziblen 3-Mannigfaltigkeiten. Außerdem wird die Aufgabe gelöst, von zwei abgeschlossenen Punktmengen in einer 3-Mannigfaltigkeit M^3 zu entscheiden, ob sie miteinander *verkettet* seien, oder ob sie sich in M^3 durch 2-Sphären trennen lassen. Ferner

---

[1]) Große lateinische Buchstaben bezeichnen stets *endliche polyederartige Punktmengen* (also Summen von endlich vielen endlichen offenen konvexen Raumstücken) in einem linearen Raum $\mathfrak{R}^n$ hinreichend hoher Dimension (s. REIDEMEISTER [12]); entsprechend können darunter auch *krummlinige Polyeder* (ALEXANDROFF-HOPF [2]) bzw. *Komplexe* im Sinne von SEIFERT-THRELFALL [13] verstanden werden (d.h. topologische Bilder von polyederartigen Punktmengen des euklidischen Raumes $\mathfrak{E}^n$).

Die gegebene 3-Mannigfaltigkeit M^3 setzen wir als *abgeschlossen* voraus, d.h. als abgeschlossen in $\mathfrak{R}^n$, bzw., wenn sie speziell ein euklidisches Polyeder oder topologisches Bild eines solchen ist, als *kompakt*. Wie MOISE [8; 9] und BING [3] bewiesen haben, sind alle 3-Mannigfaltigkeiten triangulierbar; also erfassen wir mit diesen kombinatorischen Untersuchungen auch alle kompakten 3-Mannigfaltigkeiten im Sinne der mengentheoretischen Topologie (vgl. [11]). Von allen betrachteten Punktmengen in M^3 wird vorausgesetzt, daß sie *polyederartig in* M^3 liegen, d.h. daß es eine endliche (gegebenenfalls krummlinige) Simplizialzerlegung von M^3 gibt, so daß sich die betrachteten Punktmengen als Summen von (offenen) Simplizes dieser Simplizialzerlegung darstellen lassen.

Die kombinatorische *Äquivalenzklasse* von M^3 (d.h. die Klasse aller zu M^3 verwandten Polyeder) bezeichnen wir mit $\mathfrak{M}^3$. Ist M^3 ein euklidisches Polyeder oder topologisches Bild eines solchen, so ist die Äquivalenzklasse $\mathfrak{M}^3$, wie von MOISE [8; 9] und BING [3] bewiesen wurde, identisch mit der *Homöomorphieklasse* von M^3.

[2]) $\bar{A}$ bedeutet die *abgeschlossene Hülle* der Punktmenge A (in $\mathfrak{R}^n$ bzw. in M^3), $\dot{A}$ den *Rand* ($\dot{A} = \dot{\bar{A}} + [\bar{A} - A]$, wobei der Rand $\dot{\bar{A}}$ der abgeschlossenen Punktmenge $\bar{A}$ nach [12] bzw. [2; 13] definiert ist), $\mathring{A}$ das *Innere* ($\mathring{A} = A - \dot{A}$). Die Differenz $A - B$ bedeutet die Menge der nicht in B liegenden Punkte aus A (wobei B nicht in A zu liegen braucht).

[3]) Auf Grund des von PAPAKYRIAKOPOULOS [10] und WHITEHEAD [14] bewiesenen Sphärentheorems ist für eine *irreduzible orientierbare* 3-Mannigfaltigkeit stets $\pi_2 = 0$.

wird das in [5] dargestellte Kreislinienkriterium für *beliebige* 3-Mannigfaltigkeiten verallgemeinert.

Die Ergebnisse werden mit Hilfe der Theorie der Normalflächen [5] und eines Beweisgedankens von H. KNESER [6] erhalten. Die Theorie der Normalflächen beschäftigt sich mit der Bestimmung *inkompressibler* Flächen in 3-Mannigfaltigkeiten. Als „inkompressibel in M^3" bezeichnen wir eine zusammenhängende Fläche M^2 (deren Rand entweder leer ist oder in $\dot{M}^3$ liegt, und deren Inneres in $\mathring{M}^3$ liegt) dann, wenn sie 1. keine triviale 2-Sphäre ist, d.h. keine 2-Sphäre, die ein in M^3 liegendes 3-Element berandet, und wenn sie 2. keinen trivialen Henkel besitzt, d.h. wenn es keine zweiseitig in $\mathring{M}^2$ liegende 1-Sphäre gibt, die ein in $\mathring{M}^3 - M^2$ liegendes (offenes) 2-Element berandet, jedoch kein in M^2 liegendes 2-Element berandet. (Eine nicht zusammenhängende Fläche heißt inkompressibel, wenn ihre sämtlichen Komponenten inkompressibel sind.) Wie bereits in [5] (Einleitung) angemerkt wurde, gilt für *irreduzible* 3-Mannigfaltigkeiten der Satz, daß es zu jeder inkompressiblen unberandeten Fläche eine isotope Normalfläche gibt[4]), daß man also bei irreduziblen 3-Mannigfaltigkeiten durch Betrachtung der Normalflächen *alle* inkompressiblen Flächen erfaßt. Aus diesem Grunde bedeutet es für die weitere Anwendung der Theorie der Normalflächen einen besonderen Vorteil, daß man sich auf Grund der Resultate dieser Arbeit auf die Behandlung irreduzibler 3-Mannigfaltigkeiten beschränken kann. Außerdem gilt für die inkompressiblen Flächen in irreduziblen 3-Mannigfaltigkeiten ein Endlichkeitssatz, der dem von H. KNESER in [6] bewiesenen Endlichkeitssatz für inkompressible 2-Sphären entspricht. (Beide Endlichkeitssätze werden in dieser Arbeit mit Hilfe des von H. KNESER angegebenen Beweisgedankens bewiesen.)

§ 1. Grundbegriffe und Übersicht

Diese Arbeit beschäftigt sich in erster Linie mit der Bestimmung inkompressibler 2-Sphären in einer beliebig vorgelegten 3-Mannigfaltigkeit M^3. Ist in $\mathring{M}^3$ eine 2-Sphäre S^2 gegeben, so haben wir die beiden Fälle zu unterscheiden, daß S^2 *zerlegend* oder *nichtzerlegend* in M^3 liegt. Im ersten Falle besteht $M^3 - S^2$ aus zwei zusammenhängenden Komponenten (im folgenden schreiben wir kurz z-Komponenten) Z_1^3 und Z_2^3. — Zerlegt S^2 nicht, so bezeichnen wir eine kleine Umgebung U^3 von S^2 in M^3 als einen 3-dimensionalen *Henkel* von M^3 (vgl. [10])[5]). Wir können dann $M^3 - U^3$ durch zwei 3-Elemente E_1^3 und E_2^3 (die von den beiden 2-Sphären aus $\dot{U}^3$ berandet werden und im übrigen mit M^3 punktfremd sind) zu einer 3-Mannigfaltigkeit M'^3 ergänzen (vgl. [6; 10]) und nennen diesen Prozeß eine *Henkelreduktion*. Entsprechend sagen wir, daß die Äquivalenzklasse $\mathfrak{M}'^3$ durch eine Henkelreduktion aus $\mathfrak{M}^3$

[4]) Bei geeigneter Verschärfung der Begriffe „irreduzibel" und „inkompressibel" läßt sich ein entsprechender Satz auch für berandete Flächen beweisen. Dieses wird in einer späteren Arbeit ausgeführt werden.

[5]) Modulo der Poincaréschen Vermutung ist für *orientierte* 3-Mannigfaltigkeiten „irreduzibel" äquivalent mit „henkelfrei und *indecomposable*" im Sinne von MILNOR [7].

hervorgehe. Lassen sich an M^3 successive h Henkelreduktionen durchführen, jedoch nicht $h+1$, so bezeichnen wir h als die *Henkelzahl* von M^3 und von $\mathfrak{M}^3$.

Ist in $\mathring{M}^3$ ein System von m 2-Sphären $S_1^2, \ldots, S_m^2$ gegeben (die miteinander punktfremd sind), so bezeichnen wir diese 2-Sphären als *geometrisch unabhängig in* M^3, wenn keine der z-Komponenten von $M^3-(S_1^2+\cdots+S_m^2)$ eine gelochte 3-Sphäre[6]) ist, anderenfalls als geometrisch abhängig. Entsprechend bezeichnen wir die 2-Sphäre S_i^2 (i zwischen 1 und m) als geometrisch unabhängig von den übrigen 2-Sphären, wenn diejenigen z-Komponenten von $M^3-(S_1^2+\cdots+S_m^2)$, die S_i^2 im Rande enthalten, keine gelochten 3-Sphären sind. Als *einfach* bezeichnen wir das Sphärensystem, wenn jede der 2-Sphären im Rande *zweier* z-Komponenten von $M^3-(S_1^2+\cdots+S_m^2)$ liegt. Ein einfaches Sphärensystem nennen wir *vollständig*, wenn jede in $\mathring{M}^3-(S_1^2+\cdots+S_m^2)$ liegende 2-Sphäre von $S_1^2, \ldots, S_m^2$ geometrisch abhängig ist.

$S_1^2, \ldots, S_m^2$ sei nun ein einfaches Sphärensystem in M^3. Von besonderer Bedeutung sind die z-Komponenten $Z_1^3, \ldots, Z_p^3$ von $M^3-(S_1^2+\cdots+S_m^2)$. Die im Rande von M^3 liegenden 2-Sphären seien $S_{m+1}^2, \ldots, S_{m+l}^2$ (l gegebenenfalls gleich Null). Die Zahl l bezeichnen wir als die *Randsphärenzahl* von M^3 und von $\mathfrak{M}^3$. Wir wählen $2m+l$ 3-Elemente $E_1^3, E_1^{*3}, \ldots, E_m^3, E_m^{*3}, E_{m+1}^3, \ldots, E_{m+l}^3$, deren Innenteile miteinander und mit M^3 punktfremd sind, mit $\dot{E}_i^3=\dot{E}_i^{*3}=S_i^2$. Durch diese 3-Elemente ergänzen wir $Z_1^3, \ldots, Z_p^3$ zu 3-Mannigfaltigkeiten $M_1^3, \ldots, M_p^3$. Diese 3-Mannigfaltigkeiten bezeichnen wir als *Erzeugende von* M^3 und ihre Äquivalenzklassen als Erzeugende von $\mathfrak{M}^3$. Die Gesamtheit der Erzeugenden $M_1^3, \ldots, M_p^3$ bzw. $\mathfrak{M}_1^3, \ldots, \mathfrak{M}_p^3$ nennen wir eine ($S_1^2+\cdots+S_m^2$ entsprechende) *Aufspaltung von* M^3 bzw. $\mathfrak{M}^3$ *in Erzeugende*. Ist M^3 eine orientierte 3-Mannigfaltigkeit, so erhalten wir entsprechend *orientierte Erzeugende*. Da die 3-Sphäre offenbar beliebig oft als Erzeugende einer jeden 3-Mannigfaltigkeit auftreten kann, bezeichnen wir sie als *triviale Erzeugende*. Die Erzeugenden sind dann und nur dann sämtlich irreduzibel, wenn das Sphärensystem $S_1^2, \ldots, S_m^2$ in M^3 vollständig ist. Von einem vollständigen Sphärensystem läßt sich also eine Aufspaltung in irreduzible Erzeugende ablesen. (Ferner sind die Erzeugenden dann und nur dann sämtlich nichttrivial, wenn $S_1^2, \ldots, S_m^2$ in M^3 geometrisch unabhängig sind.)

In dieser Arbeit wird ein Verfahren angegeben, in einer beschränkten Anzahl von Schritten die Henkelzahl einer beliebig gegebenen 3-Mannigfaltigkeit M^3 zu bestimmen und ein vollständiges Sphärensystem zu konstruieren.

Wie H. Kneser [6] bemerkt hat, sind die nichttrivialen irreduziblen Erzeugenden und die Henkelzahl von $\mathfrak{M}^3$ eindeutig festgelegt, und umgekehrt ist $\mathfrak{M}^3$ durch die irreduziblen Erzeugenden, die Henkelzahl, die Randsphärenzahl und den Orientierbarkeitscharakter eindeutig bestimmt. Ein Beweis dieses Sachverhaltes ergibt sich in §3 dieser Arbeit. Ungelöst bleibt nach wie vor die Aufgabe, von einer Erzeugenden zu entscheiden, ob sie trivial sei,

---

[6]) Unter einer „gelochten 3-Sphäre in M^3" verstehen wir nach [6] eine Punktmenge, deren abgeschlossene Hülle in M^3 ein von 2-Sphären berandeter Teil einer 3-Sphäre ist.

d.h. ein notwendiges und hinreichendes Äquivalenzkriterium für die 3-Sphäre anzugeben.

Von H. KNESER [6] wurde bewiesen, daß es in einer gegebenen 3-Mannigfaltigkeit M^3 nur eine beschränkte Anzahl von geometrisch unabhängigen 2-Sphären geben kann. Dieser Endlichkeitssatz ist in [6] nur für unberandete 3-Mannigfaltigkeiten formuliert. Der Beweis gilt aber ebenso für berandete 3-Mannigfaltigkeiten und läßt sich ferner für beliebige inkompressible Flächen in irreduziblen 3-Mannigfaltigkeiten erweitern, wie wir in §3 zeigen werden. Mit dem Kneserschen Endlichkeitssatz ist sichergestellt, daß ein vollständiges Sphärensystem in M^3 *existiert*. Auf dieses Resultat aufbauend wird im folgenden ein Verfahren entwickelt, nach dem man ein solches vollständiges Sphärensystem (in einer beschränkten Anzahl von Schritten) *konstruieren* kann. Die hierfür verwendete Normalflächenmethode wurde bereits in [5] dargestellt.

Zu einer beliebigen Zerlegung Γ von M^3 in offene Raumelemente wird eine Normalzerlegung konstruiert, die Normalflächen bezüglich dieser Normalzerlegung werden in Ähnlichkeitsklassen eingeteilt und es werden unter diesen die endlich vielen, eindeutig ausgezeichneten, Fundamentalklassen bestimmt. In dieser Arbeit (§2) werden zunächst einige Sätze über die sog. ξ-Operationen bewiesen, durch die sich nach [5] jede Fläche in eine Normalfläche überführen läßt. §3 enthält die bereits erwähnten Endlichkeits- und Eindeutigkeitssätze. In §4 wird dann gezeigt: Gibt es eine einseitig in $\mathring{M}^3$ liegende projektive Ebene, so sind die Repräsentanten mindestens einer der Fundamentalklassen ebenfalls solche projektiven Ebenen. Eine kleine Umgebung einer derartigen projektiven Ebene wird von einer 2-Sphäre berandet und wir bezeichnen sie als eine 3-dimensionale *Kreuzhaube*; sie kann durch ein 3-Element zum projektiven Raum ergänzt werden. Wir können in M^3 die Kreuzhaube durch ein 3-Element mit gleichem Rand (dessen Inneres mit M^3 punktfremd ist) ersetzen und nennen diesen Prozeß eine *Kreuzhaubenreduktion*. Dieses Verfahren wenden wir so oft an, bis wir eine 3-Mannigfaltigkeit M'^3 erhalten, in der es keine einseitige projektive Ebene mehr gibt. Wurde M'^3 durch k Kreuzhaubenreduktionen erhalten, so nennen wir k die *Kreuzhaubenzahl* von M^3 und von $\mathfrak{M}^3$; $\mathfrak{M}^3$ enthält dann den projektiven Raum k mal als (irreduzible) Erzeugende.

Für die kreuzhaubenfreie 3-Mannigfaltigkeit M'^3 beweisen wir dann weiter: Gibt es in $\mathring{M}'^3$ eine nichtzerlegende 2-Sphäre, so sind die Repräsentanten mindestens einer Fundamentalklasse ebenfalls nichtzerlegende 2-Sphären. Diese Beweise werden nach demselben Prinzip geführt, das bereits in [5] verwendet wurde (s. Abschnitt d der Einleitung von [5]). Gibt es eine nichtzerlegende 2-Sphäre in $\mathring{M}'^3$, so können wir eine Henkelreduktion durchführen und dieses Verfahren so oft wiederholen, bis nach h Schritten eine henkel- und kreuzhaubenfreie 3-Mannigfaltigkeit M''^3 entsteht.

Mit diesen Resultaten lassen sich das in [5] angegebene Bestimmungsverfahren für Bänder minimaler Charakteristik und das daraus folgende Kreis-

linienkriterium in einfacher Weise so verallgemeinern, daß sie für beliebige 3-Mannigfaltigkeiten gelten: In der vorgelegten 3-Mannigfaltigkeit M^3 werden zunächst alle Kreuzhauben und Henkel reduziert. Die dadurch entstehende 3-Mannigfaltigkeit M'^3 erfüllt die in [5] gemachten Voraussetzungen, so daß das Bestimmungsverfahren angewendet werden kann. Wie sich leicht zeigen läßt, wird die minimale Charakteristik der Bänder, die sich in eine gegebene, in M^3 verlaufende Linie einspannen lassen, durch die Kreuzhauben- und Henkelreduktionen nicht geändert. Also ist das in M'^3 konstruierte Band minimaler Charakteristik auch von minimaler Charakteristik in M^3.

In §5 können wir bereits eine henkel- und kreuzhaubenfreie 3-Mannigfaltigkeit voraussetzen. Es ließe sich nun leicht beweisen: Gibt es in dieser 3-Mannigfaltigkeit M^3 eine inkompressible 2-Sphäre, so sind auch die Repräsentanten mindestens einer der Fundamentalklassen inkompressible 2-Sphären. Wenn wir entscheiden könnten, ob eine vorgelegte 2-Sphäre in M^3 inkompressibel sei, so könnten wir damit feststellen, ob es in M^3 eine inkompressible 2-Sphäre gebe und M^3 bei positivem Ergebnis in zwei nichttriviale Erzeugende aufspalten; für diese Erzeugenden könnte das Verfahren wiederholt werden, bis schließlich die (nichttrivialen) irreduziblen Erzeugenden erhalten würden. Da aber bisher kein notwendiges und hinreichendes Äquivalenzkriterium für das 3-Element (bzw. für die 3-Sphäre) bekannt ist, können wir im allgemeinen nicht entscheiden, ob eine 2-Sphäre in M^3 kompressibel sei. Wir können daher auf diesem Wege nicht in jedem Falle zu einem vollständigen Sphärensystem gelangen.

Mit etwas größerem Aufwand läßt sich jedoch folgendes beweisen: Es existiert in M^3 ein vollständiges Sphärensystem, so daß jede der 2-Sphären in einer der Fundamentalklassen liegt[7]). Damit können wir, von den Fundamentalklassen ausgehend, ein Sphärensystem in M^3 konstruieren, das mit Sicherheit vollständig ist (jedoch im allgemeinen nicht geometrisch unabhängig).

Das Bestimmungsverfahren für ein vollständiges Sphärensystem gestattet insbesondere die Entscheidung, ob s gegebene (miteinander punktfremde) abgeschlossene Punktmengen $P_1, \ldots, P_s$ in einer beliebigen 3-Mannigfaltigkeit M^3 miteinander verkettet seien; wir bezeichnen $P_1, \ldots, P_s$ als *miteinander verkettet in* M^3, wenn folgendes gilt: Sind $S_1^2, \ldots, S_m^2$ beliebige, miteinander und mit $P_1, \ldots, P_s$ punktfremde 2-Sphären in M^3, so liegen $P_1, \ldots, P_s$ stets in einer und derselben z-Komponente von $M^3 - (S_1^2 + \cdots + S_m^2)$[8]).

[7]) Zu dem in dieser Arbeit gegebenen Beweis bemerkte H. SCHUBERT, daß sich bei anderer Fassung die Voraussetzung der Henkelfreiheit von M^3 vermeiden läßt und zugleich ein anderer Beweis des Kneserschen Endlichkeitssatzes erhalten werden kann.

[8]) „verkettet" ist in diesem Sinne das Gegenteil von *geometrical splittable* (s. [10]). Ist M^3 *henkelfrei*, so braucht statt des Sphärensystems $S_1^2, \ldots, S_m^2$ nur eine einzelne 2-Sphäre betrachtet zu werden. Denn in diesem Falle ist jede 2-Sphäre in M^3 zerlegend und es gilt: Lassen sich $P_1, \ldots, P_s$ in M^3 durch ein Sphärensystem trennen, so werden sie bereits durch *eine* (geeignete) der 2-Sphären getrennt.

§ 2. Sätze über ξ-Operationen

Ist M^3 eine 3-Mannigfaltigkeit, Γ eine beliebige randtreue Zerlegung von M^3 in offene Raumelemente und Θ eine Normalzerlegung bezüglich Γ, so läßt sich, wie in [5], Kap. I ausgeführt wurde, jede in M^3 liegende 2-Mannigfaltigkeit M^2 durch sog. ξ-Operationen in eine Normalfläche bezüglich Θ überführen (Normalflächensatz). Eine ξ-Operation besteht darin, daß der Durchschnitt von M^2 mit einem (in M^3 liegenden) 3-Element E^3 fortgelassen wird und in die entstandenen Lochränder (in E^3 liegende) 2-Elemente eingesetzt werden; das 3-Element E^3 wollen wir als den *Kern* der ξ-Operation bezeichnen. Für den Normalflächensatz wurde in [5] vorausgesetzt, daß M^2 1. „transversal" in M^3 liegt, d.h. daß die Randlinien von M^2 (falls M^2 berandet ist) in $\dot{M}^3$ verlaufen und $\mathring{M}^2$ in $\mathring{M}^3$ liegt, und daß sich 2. die Randlinien von M^2 aus 0- und 1-Elementen aus Γ zusammensetzen lassen. (Daß sich die Voraussetzung 2 bei geeigneter Verallgemeinerung des Begriffes „ξ-Operation" vermeiden läßt, wurde bereits in [5] angemerkt, soll aber erst in einer späteren Arbeit ausgeführt werden.) In diesem Paragraphen beweisen wir eine Reihe von Sätzen, die zum Teil bereits in [5] angekündigt wurden und besagen, daß gewisse für M^3 charakteristische Eigenschaften einer Fläche durch ξ-Operationen nicht verloren gehen. Die beiden wichtigsten dieser Sätze besagen: 1. Ist M^3 irreduzibel und M^2 eine inkompressible Fläche, so entsteht durch ξ-Operationen stets (unter anderem) eine zu M^2 isotope Fläche; 2. Ist M^3 beliebig und M^2 ein vollständiges Sphärensystem, so geht M^2 durch ξ-Operationen stets in ein vollständiges Sphärensystem über. Daraus und aus dem Normalflächensatz ergibt sich u.a. die Darstellbarkeit der entsprechenden Flächen als Normalflächen.

Zunächst bemerken wir, daß sich jede ξ-Operation durch eine endliche Folge von einfacheren Operationen ersetzen läßt: Ist E^3 ein 3-Element in M^3, das einen Kreisring K^2 aus M^2 im Rande enthält und im übrigen mit M^2 punktfremd ist (ist also $M^2 E^3 = K^2$, wenn wir den Durchschnitt zweier Punktmengen A und B der Einfachheit halber als Produkt AB schreiben), so nennen wir die Ersetzung von $\mathring{K}^2$ durch $\dot{E}^3 - K^2$ eine an M^2 ausführbare *ϱ-Operation in* M^3 und das 3-Element E^3 den *Kern* der ϱ-Operation. — Ist E'^3 ein 3-Element in M^3 mit $M^2 E'^3 = \dot{E}'^3$ (ist also die Randsphäre $\dot{E}'^3$ von E'^3 eine z-Komponente von M^2), so nennen wir die Fortlassung von $\dot{E}'^3$ eine an M^2 ausführbare *Sphärenreduktion in* M^3 und E'^3 den Kern derselben.

SATZ 1. *In einer beliebigen 3-Mannigfaltigkeit M^3 gilt: Geht M' aus einer 2-Mannigfaltigkeit M^2 durch eine ξ-Operation hervor, so läßt sich M^2 auch durch eine endliche Folge von ϱ-Operationen und Sphärenreduktionen (und isotopen Deformationen) in M' überführen. (Ist E^3 der Kern der ξ-Operation, so lassen sich alle Operationen innerhalb von E^3 durchführen.)*

BEWEIS. E^3 sei der Kern der ξ-Operation; die Schnittlinien von M^2 mit $\dot{E}^3$ seien $L_1^1, \ldots, L_s^1$ (dabei kann s gleich Null sein). In E^3 betrachten wir s miteinander punktfremde 2-Elemente $E_1^2, \ldots, E_s^2$, die von $L_1^1, \ldots, L_s^1$ berandet

werden und im übrigen mit $\dot{E}^3$ punktfremd sind. Diese 2-Elemente wählen wir so, daß sie mit M^2 außer ihren Rändern höchstens noch unberandete Durchdringungslinien gemeinsam haben; (dieses ist stets möglich, da sich etwaige Berührungspunkte, -Kanten oder -Flächenstücke durch kleine isotope Deformationen beseitigen lassen). Sind Durchdringungslinien vorhanden, so zerlegen diese die $E_1^2, \ldots, E_s^2$ in z-Komponenten, deren mindestens eine ein 2-Element E^{*2} ist. Dann wählen wir in M^2 eine kleine Umgebung U^2 von $\dot{E}^{*2}$; U^2 ist ein Kreisring. In die beiden Randlinien von U^2 spannen wir zwei nahe zu E^{*2} benachbarte 2-Elemente E'^2, E''^2 ein (s. Fig. 1)[9]. Die Ersetzung von U^2 durch $E'^2 + E''^2$ ist eine ϱ-Operation, durch die M^2 in eine 2-Mannigfaltigkeit übergeht, die eine Durchdringungslinie mit den $E_1^2, \ldots, E_s^2$ weniger hat als M^2. So verfahren wir weiter, bis wir eine 2-Mannigfaltigkeit M_1^2 erhalten, die mit $\mathring{E}_1^2, \ldots, \mathring{E}_s^2$ punktfremd ist.

Fig. 1

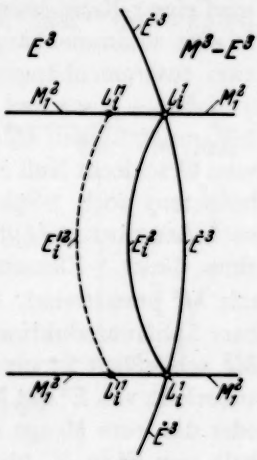

Fig. 2

In $M_1^2 E^3$ wählen wir eine nahe zu L_i^1 $(i = 1, \ldots, s)$ benachbarte Linie $L_i'^1$ und spannen in diese ein nahe zu E_i^2 benachbartes 2-Element $E_i'^2$ ein (s. Fig. 2). Dann ist die Ersetzung des zwischen L_i^1 und $L_i'^1$ liegenden Kreisringes aus M_1^2 durch die 2-Elemente E_i^2 und $E_i'^2$ eine ϱ-Operation. Derartige ϱ-Operationen führen wir für alle $i = 1, \ldots, s$ an M_1^2 aus und erhalten dadurch eine 2-Mannigfaltigkeit M_2^2.

Die z-Komponenten von $M_2^2 E^3$ sind die 2-Elemente $E_1^2, \ldots, E_s^2$ und gegebenenfalls endlich viele orientierbare unberandete 2-Mannigfaltigkeiten. Die Quadratsumme der Geschlechter dieser z-Komponenten sei q. Ist $q > 0$, so befindet sich unter den unberandeten z-Komponenten von $M_2^2 E^3$ — ihre Summe werde mit M^{*2} bezeichnet — mindestens eine, die keine 2-Sphäre

[9]) *Anmerkung zu den Figuren.* Die dieser Arbeit beigegebenen Figuren sollen die Berandungsverhältnisse der im Text behandelten Punktmengen veranschaulichen und sind folgendermaßen zu verstehen:

Im allgemeinen werden (berandete oder unberandete) 2-Mannigfaltigkeiten dargestellt, die gemeinsame (unberandete) Rand- bzw. Durchdringungslinien besitzen und in einer 3-Mannigfaltigkeit M^3 liegen. Zur Darstellung einer Rand- oder Durchdringungslinie L^1 werden ein oder zwei Punkte gezeichnet; liegt L^1 im Rande einer dargestellten 2-Mannigfaltigkeit M^2, so wird von jedem der L^1 darstellenden Punkte aus eine mit „M^2" bezeichnete Linie gezeichnet; liegt L^1 im Innern von M^2, so werden von jedem Punkte „L^1" aus zwei Linien „M^2" gezeichnet (und zwar so, daß bei einer Durchdringung von M^2 mit einer 2-Mannigfaltigkeit M'^2 auch die darstellenden Linien einander in der Zeichenebene durchdringen). Eine M^2 darstellende Linie kann gegebenenfalls zwei Punkte miteinander verbinden, die geschlossene Linien darstellen. — Entsprechend können Gebiete der Zeichenebene zur Darstellung 3-dimensionaler Teilmannigfaltigkeiten von M^3 dienen.

ist. Dann gibt es in M^{*2} nach FOX [4] eine 1-Sphäre L^{*1}, die kein in M^{*2} liegendes 2-Element berandet, jedoch ein in $\mathring{E}^3 - M^{*2}$ liegendes (offenes) 2-Element E^{*2} berandet. Dieses 2-Element E^{*2} kann so gewählt werden, daß es mit den 2-Elementen $E_1^2, \ldots, E_s^2$ punktfremd ist. Dann wählen wir in M_2^2 eine zu L^{*1} nahe benachbarte Linie $L^{*'1}$ und ein in $L^{*'1}$ eingespanntes, nahe zu E^{*2} benachbartes 2-Element $E^{*'2}$ in E^3 und ersetzen den zwischen L^{*1} und $L^{*'1}$ in M_2^2 liegenden Kreisring durch $E^{*2} + E^{*'2}$. Durch diese ϱ-Operation geht $M_2^2 E^3$ in eine 2-Mannigfaltigkeit über, deren z-Komponenten eine Quadratsumme der Geschlechter haben, die kleiner als q ist. (Denn es wird eine z-Komponente von $M_2^2 E^3$ mit Geschlecht $g > 0$ übergeführt entweder in eine zusammenhängende 2-Mannigfaltigkeit kleineren Geschlechts oder in zwei zusammenhängende 2-Mannigfaltigkeiten mit Geschlechtern g_1 und g_2, so daß $g_1 + g_2 = g$ und $g_1 \neq 0$, $g_2 \neq 0$ ist.) So fahren wir fort, bis wir zu einer 2-Mannigfaltigkeit M_3^2 gelangen, deren Durchschnitt mit E^3 nur z-Komponenten vom Geschlecht Null enthält. $M_3^2 E^3$ enthält außer den 2-Elementen $E_1^2, \ldots, E_s^2$ höchstens noch 2-Sphären. Sind solche 2-Sphären vorhanden, so berandet nach ALEXANDER [1] jede von ihnen ein in E^3 liegendes 3-Element; mindestens eines dieser 3-Elemente — es werde mit E^{*3} bezeichnet — ist im Innern mit M_3^2 punktfremd. Dann ist die Fortlassung von $\dot{E}^{*3}$ eine an M_3^2 ausführbare Sphärenreduktion. Durch eine Folge derartiger Sphärenreduktionen geht M_3^2 schließlich in eine Punktmenge M_4 mit $M_4 E^3 = E_1^2 + \cdots + E_s^2$ über, die außerhalb von E^3 mit M^2 übereinstimmt (und entweder eine 2-Mannigfaltigkeit oder die leere Menge ist). M_4 kann nun durch isotope Deformationen innerhalb von E^3 in M' überführt werden. Damit ist Satz 1 bewiesen.

SATZ 2. *Geht M'^2 aus einer 2-Mannigfaltigkeit M^2 durch eine endliche Folge von ξ-Operationen in M^3 hervor, so gilt:*

a) *Ist M^3 irreduzibel, und ist M^2 eine inkompressibel in M^3 liegende 2-Mannigfaltigkeit, so ist eine Komponente von M'^2 in M^3 isotop zu M^2.*

Für beliebige 3-Mannigfaltigkeit M^3 gilt weiter:

b) *Ist M^2 eine einseitig in $\mathring{M}^3$ liegende projektive Ebene, so ist eine z-Komponente von M'^2 ebenfalls eine einseitige projektive Ebene.*

c) *Ist M^2 eine nichtzerlegende 2-Sphäre in $\mathring{M}^3$, so ist (mindestens) eine z-Komponente von M'^2 ebenfalls eine nichtzerlegende 2-Sphäre.*

d) *Ist M^2 eine inkompressible 2-Sphäre, so existiert (mindestens) eine z-Komponente M''^2 von M'^2, die ebenfalls eine inkompressible 2-Sphäre ist.*

e) *Ist M^2 ein geometrisch unabhängiges Sphärensystem in M^3, so existiert eine Komponente M''^2 von M'^2, die ebenfalls ein geometrisch unabhängiges Sphärensystem ist und aus ebensovielen z-Komponenten besteht, wie M^2.*

BEWEIS. Auf Grund von Satz 1 genügt es, den Beweis für den Fall zu führen, daß M'^2 aus M^2 durch eine einzelne ϱ-Operation oder Sphärenreduktion hervorgeht. Nun läßt sich an M^2 überhaupt keine Sphärenreduktion durchführen. Wir brauchen daher nur noch die Auswirkung einer ϱ-Operation zu betrachten.

E^3 sei also der Kern einer ϱ-Operation, durch die M^2 in M'^2 übergeht. Der Kreisring $M^2\dot{E}^3$ werde mit K^2 bezeichnet, die beiden z-Komponenten von $\dot{E}^3-\mathring{K}^2$ mit E'^2 und E''^2.

Beweis von a). Ist M^2 inkompressibel in M^3, so beranden $\dot{E}'^2$ und $\dot{E}''^2$ in M^2 liegende 2-Elemente. Mindestens eines dieser 2-Elemente — es werde mit E^{*2} bezeichnet — ist mit $\mathring{K}^2$ punktfremd (s. Fig. 3). Bei geeigneter Wahl der Bezeichnung ist $\dot{E}^{*2}=\dot{E}'^2$. Dann ist $E^{*2}+E'^2$ eine 2-Sphäre. Ist M^3 irreduzibel, so berandet diese 2-Sphäre ein 3-Element E^{*3} in M^3. Dabei ist $\mathring{E}^{*3}$ mit M^2 punktfremd (da nach [1] und [4] jede in einem 3-Element liegende Fläche kompressibel ist, also keine z-Komponente von M^2 in E^{*3} liegen kann).

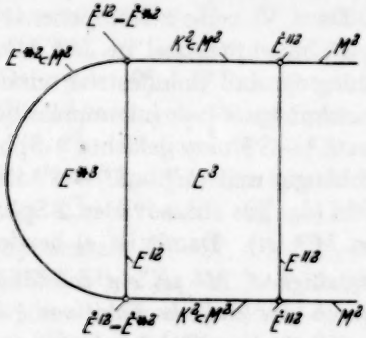

Fig. 3

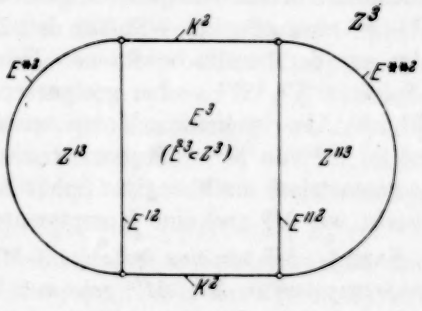

Fig. 4

Damit ist auch $\mathring{E}^{*3}+\mathring{E}'^2+\mathring{E}^3$ ein mit M^2 punktfremdes offenes 3-Element und wird von $K^2+E^{*2}+E''^2$ berandet; also ist $[M^2-(K^2+E^{*2})]+E''^2$ eine zu M^2 isotope z-Komponente von M'^2, w.z.b.w.

Beweis von b). Ist M^2 eine einseitig in M^3 liegende projektive Ebene, so besteht $M^2-\mathring{K}^2$ aus einem 2-Element E^{*2} und einem einseitig in M^3 liegenden Möbiusschen Band M^{*2}. Bei geeignet gewählter Bezeichnung ist $\dot{M}^{*2}=\dot{E}''^2$. Dann ist E''^2+M^{*2} eine einseitig in M^3 liegende projektive Ebene (und eine z-Komponente von M'^2), w.z.b.w.

Beweis von c) und d). Ist M^2 eine 2-Sphäre in $\dot{M}^3$, so besteht $M^2-\mathring{K}^2$ aus zwei 2-Elementen E^{*2} und E^{**2}, so daß (bei geeigneter Bezeichnung) $\dot{E}^{*2}=\dot{E}'^2$ und $\dot{E}^{**2}=\dot{E}''^2$ ist (s. Fig. 4). M'^2 besteht dann aus den beiden 2-Sphären $S'^2=E'^2+E^{*2}$ und $S''^2=E''^2+E^{**2}$. Damit gilt:

1. Liegen S'^2 und S''^2 zerlegend in M^3, so besteht $M^3-M'^2$ aus drei z-Komponenten Z^3, Z'^3 und Z''^3 (vgl. Fig. 4), so daß (bei geeigneter Bezeichnung) $\mathring{E}^3$ in Z^3 liegt, S'^2 in $\dot{Z}'^3$ und S''^2 in $\dot{Z}''^3$. Dann sind $\bar{Z^3}-E^3$ und $(Z'^3+Z''^3+E^3)^\circ$ die z-Komponenten von M^3-M^2, also liegt dann auch M^2 zerlegend in M^3. Daraus folgt Behauptung c).

2. Ist (mindestens) eine z-Komponente von $M^3-S'^2$ und (mindestens) eine z-Komponente von $M^3-S''^2$ ein (offenes) 3-Element, so besteht $M^3-M'^2$

wieder aus drei z-Komponenten Z^3, Z'^3 und Z''^3 (s. Nr. 1 und Fig. 4) und mindestens zwei von diesen sind gelochte 3-Sphären. Daraus folgt, daß mindestens eine der beiden z-Komponenten von $M^3 - M^2$ ein 3-Element ist, also auch M^2 nicht inkompressibel. Daraus folgt Behauptung d).

Beweis von e). Ist M^2 ein geometrisch unabhängiges Sphärensystem in M^3, so liegt der Kreisring K^2 in einer 2-Sphäre S^2 aus M^2. Weiter liegt S^2 in einer z-Komponente von $M^3 - (M^2 - S^2)$, die wir mit K^3 bezeichnen. M'^2 entsteht aus M^2 dadurch, daß S^2 durch die ϱ-Operation in zwei 2-Sphären S'^2 und S''^2 übergeht. Die z-Komponenten von $M^3 - M'^2$ sind die mit K^3 punktfremden z-Komponenten von $M^3 - M^2$ und die z-Komponenten von $K^3 - (S'^2 + S''^2)$. — Wir ergänzen K^3 durch 3-Elemente (die miteinander punktfremd sind, und deren Innenteile mit M^3 punktfremd sind) zu einer randsphärenfreien 3-Mannigfaltigkeit M'^3. Da n. V. keine z-Komponente von $K^3 - S^2$ eine gelochte 3-Sphäre ist, liegt S^2 inkompressibel in M'^3. Damit folgt aus der bereits bewiesenen Behauptung d), daß (mindestens) eine der 2-Sphären S'^2, S''^2 — bei geeigneter Bezeichnung S'^2 — inkompressibel in M'^3 ist. Also ist keine z-Komponente von $K^3 - S'^2$ eine gelochte 3-Sphäre; also ist S'^2 von $M^2 - S^2$ geometrisch unabhängig und $(M^2 - S^2) + S'^2$ damit ein geometrisch unabhängiges Sphärensystem (das aus ebensovielen 2-Sphären besteht wie M^2 und eine Komponente von M'^2 ist). Damit ist e) bewiesen.

SATZ 3. *M^3 sei eine beliebige 3-Mannigfaltigkeit, M^2 sei ein vollständiges Sphärensystem in M^3. M'^2 gehe aus M^2 durch eine endliche Folge von ξ-Operationen hervor. Dann ist M'^2 ebenfalls ein vollständiges Sphärensystem in M^3 und es gilt:*

a) *Sind m und m' die Anzahlen der Sphären von M^2 bzw. M'^2, p und p' die Anzahlen der z-Komponenten von $M^3 - M^2$ bzw. $M^3 - M'^2$, so ist $m - p = m' - p'$.*

b) *Sind $(M_1^3, \ldots, M_p^3)$ und $(M_1'^3, \ldots, M_{p'}'^3)$ zwei M^2 bzw. M'^2 entsprechende Aufspaltungen von M^3 in Erzeugende und sind dabei $M_1^3, \ldots, M_q^3$ nichttrivial und $M_{q+1}^3, \ldots, M_p^3$ trivial, so ist bei geeigneter Numerierung M_i^3 kombinatorisch äquivalent zu $M_i'^3$ für alle $i = 1, \ldots, q$ und $M_{q+1}'^3, \ldots, M_{p'}'^3$ sind trivial.*

Ist M^3 eine orientierte 3-Mannigfaltigkeit, so sind auch die entsprechend orientierten Erzeugenden $M_1^3, \ldots, M_q^3$ und $M_1'^3, \ldots, M_q'^3$ paarweise äquivalent.

BEWEIS. Die 2-Sphären, aus denen M^2 besteht, seien $S_1^2, \ldots, S_m^2$; die z-Komponenten von $M^3 - (S_1^2 + \cdots + S_m^2)$ seien $Z_1^3, \ldots, Z_p^3$. — Nach Satz 1 brauchen nur die beiden Fälle behandelt zu werden, daß M'^2 durch eine einzelne ϱ-Operation aus M^2 hervorgeht und daß M'^2 durch eine einzelne Sphärenreduktion aus M^2 entsteht.

1. Beweis für den Fall einer Sphärenreduktion. Geht M'^2 aus M^2 durch eine Sphärenreduktion hervor, so wird bei geeigneter Numerierung die 2-Sphäre S_m^2 aus M^2 fortgelassen und die z-Komponente Z_p^3 ist ein offenes 3-Element mit $Z_p^3 = S_m^2$; außerdem liegt S_m^2 im Rande von Z_{p-1}^3. Dann besteht M'^2 aus den 2-Sphären $S_1'^2 = S_1^2, \ldots, S_{m-1}'^2 = S_{m-1}^2$, und die z-Komponenten von $M^3 - M'^2$ sind $Z_1'^3 = Z_1^3, \ldots, Z_{p-2}'^3 = Z_{p-2}^3$, $Z_{p-1}'^3 = Z_{p-1}^3 + S_m^2 + Z_p^3$. Also ist

M'^3 ein einfaches Sphärensystem in M^3, es gilt die Behauptung a) ($m'=m-1$; $p'=p-1$) und bei Ergänzung von $Z_1^3, \ldots, Z_p^3$ und von $Z_1'^3, \ldots, Z_{p-1}'^3$ zu Erzeugenden $M_1^3, \ldots, M_p^3$ bzw. $M_1'^3, \ldots, M_{p-1}'^3$ ist $M_i'^3$ äquivalent zu M_i^3 für alle $i=1, \ldots, p-1$ und M_p^3 ist trivial. Also gilt auch Behauptung b), und Satz 3 ist damit für den Fall einer Sphärenreduktion bewiesen.

2. Beweis für den Fall einer ϱ-Operation. Geht M'^3 aus M^3 durch eine ϱ-Operation — ihr Kern werde mit E^3 bezeichnet — hervor, so gilt bei geeigneter Numerierung: $\mathring{E}^3$ liegt in Z_p^3; der in M^3 liegende Kreisring K^2 aus $\dot{E}^3$ liegt in S_m^2; S_m^2 liegt in $\dot{Z}_{p-1}^3$ und in $\dot{Z}_p^3$. $\dot{E}^3-\mathring{K}^2$ besteht aus zwei 2-Elementen E'^2 und E''^2, $S_m^2-\mathring{K}^2$ besteht aus zwei 2-Elementen E^{*2} und E^{**2}, so daß $\dot{E}'^2=\dot{E}^{*2}$ und $\dot{E}''^2=\dot{E}^{**2}$ ist (vgl. Fig. 4). Damit gilt bei geeignet gewählter Bezeichnung weiter: $Z_p^3-E^3$ besteht aus zwei z-Komponenten Z'^3 und Z''^3, so daß E'^2 in $\dot{Z}'^3$ und E''^2 in $\dot{Z}''^3$ liegt und so daß Z''^3 eine gelochte 3-Sphäre ist. (Denn wäre $Z_p^3-E^3$ zusammenhängend, oder wären beide z-Komponenten von $Z_p^3-E^3$ keine gelochten 3-Sphären, so wäre eine nahe zu E'^2+E^{*2} benachbarte, in $Z_p^3-E^3$ liegende 2-Sphäre geometrisch unabhängig von M^2, also M^2 nicht vollständig.) Also besteht M'^3 aus den 2-Sphären $S_1'^2=S_1^2, \ldots, S_{m-1}'^2=S_{m-1}^2$, $S_m'^2=E^{*2}+E'^2$, $S_{m+1}'^2=E^{**2}+E''^2$, und $M^3-M'^3$ besteht aus den z-Komponenten $Z_1'^3=Z_1^3, \ldots, Z_{p-2}'^3=Z_{p-2}^3$, $Z_{p-1}'^3=Z_{p-1}^3+\mathring{K}^3+\mathring{E}^3$, $Z_p'^3=Z'^3=Z_p^3-(Z''^3+E^3)$, $Z_{p+1}'^3=Z''^3$. Also ist M'^3 ein einfaches Sphärensystem und es gilt die Behauptung a) ($m'=m+1$, $p'=p+1$). Werden $Z_1^3, \ldots, Z_p^3$ und $Z_1'^3, \ldots, Z_{p+1}'^3$ durch 3-Elemente zu Erzeugenden $M_1^3, \ldots, M_p^3$ bzw. $M_1'^3, \ldots, M_{p+1}'^3$ ergänzt, so ist offenbar $M_i'^3$ kombinatorisch äquivalent zu M_i^3 für alle $i=1, \ldots, p-2$ und $M_{p+1}'^3$ ist trivial; wie sich leicht zeigen läßt, ist auch $M_{p-1}'^3$ äquivalent zu M_{p-1}^3 und $M_p'^3$ äquivalent zu M_p^3: Ist $E_1'^3$ dasjenige der zur Ergänzung von $Z_{p-1}'^3$ eingesetzten 3-Elemente, dessen Rand E'^2+E^{*2} ist, so ist $\overline{Z}_{p-1}'^3+E_1'^3$ äquivalent zu $\overline{Z}_{p-1}^3$, also ist auch $M_{p-1}'^3$ äquivalent zu M_{p-1}^3; sind $E_1^3, \ldots, E_s^3$ diejenigen der zur Ergänzung von Z_p^3 eingesetzten 3-Elemente, deren Ränder in $\dot{Z}''^3$ liegen, so ist $\overline{Z}_p'^3$ äquivalent zu $\overline{Z}_p^3+E_1^3+\cdots+E_s^3$, also ist auch $M_p'^3$ äquivalent zu M_p^3. Damit ist auch Behauptung b) bewiesen und der Beweis von Satz 3 für den Fall einer ϱ-Operation vollendet.

Mit Nr. 1 und Nr. 2 ist Satz 3 bewiesen.

§ 3. Endlichkeits- und Eindeutigkeitssätze

Mit Hilfe der Sätze 2 a) und 2 e) des vorigen Paragraphen und des Hauptsatzes 1 aus [5] können wir nun zwei Endlichkeitssätze beweisen, die besagen, daß es in einer 3-Mannigfaltigkeit M^3 nur eine beschränkte Anzahl miteinander punktfremder inkompressibler Flächen geben kann, ohne daß mindestens zwei von ihnen topologisch parallel zueinander sind. *Topologisch parallel in* M^3 nennen wir zwei Flächen M_1^2, M_2^2 dann, wenn es eine z-Komponente M'^3 von $M^3-(M_1^2+M_2^2)$ gibt, über die sich M_1^2 isotop in M_2^2 verschieben läßt. Von H. Kneser wurde in [6] eine Methode angegeben, nach der man Sätze dieser Art beweisen kann: Man betrachtet ein beliebiges System

von miteinander punktfremden Flächen (in [6] 2-Sphären) in M^3 und eine beliebig gewählte Zerlegung Ω (in [6] eine Simplizialzerlegung) von M^3; dann ändert man das Flächensystem so ab, daß es einerseits die Elemente aus Ω in einer einfachen Weise schneidet und daß andererseits seine wesentlichen topologischen Eigenschaften erhalten bleiben; schließlich betrachtet man die durch das abgeänderte Flächensystem bewirkte Unterteilung von Ω und stellt fest, daß sie nur eine beschränkte Anzahl von 3-dimensionalen Elementen enthalten kann, die nicht „Prismen" sind; daraus folgt das Weitere. — Wir verwenden als Zerlegung Ω eine Normalzerlegung Θ von M^3 und als Normalform für das Flächensystem die Normalflächeneigenschaft. Ist Q^3 ein P-, K- oder F-Raumstück aus Θ, und sind Q_1^2 und Q_2^2 zwei miteinander punktfremde und bezüglich Θ zueinander ähnliche Normalflächenstücke in $\overline{Q}^3$, so bezeichnen wir die zwischen Q_1^2 und Q_2^2 liegende z-Komponente von $Q^3-(Q_1^2+Q_2^2)$ als ein *P-*, *K-*, bzw. *F-Normalprisma (bezüglich Θ)*. Dann beweisen wir einen Endlichkeitssatz über Normalflächen, aus dem sich [zusammen mit Satz 2 a), 2 e) und dem Normalflächensatz] die Endlichkeitssätze für inkompressible Flächen ergeben. Aus den Endlichkeitssätzen ergibt sich unmittelbar die *Existenz* eines vollständigen Sphärensystems in jeder beliebigen 3-Mannigfaltigkeit M^3 (s. [6]). Mit Hilfe des Satzes 3 läßt sich außerdem leicht die *Eindeutigkeit* der Aufspaltung von $\mathfrak{M}^3$ in irreduzible Erzeugende (bis auf triviale Erzeugende) beweisen. Das heißt: Sind zwei beliebige vollständige Sphärensysteme in M^3 gegeben, so sind die *nichttrivialen* Erzeugenden der entsprechenden Aufspaltungen von M^3 paarweise kombinatorisch äquivalent; ferner läßt sich die Henkelzahl von M^3 (und $\mathfrak{M}^3$) von jedem beliebigen vollständigen Sphärensystem ablesen als Anzahl der 2-Sphären minus Anzahl der in der entsprechenden Aufspaltung vorkommenden Erzeugenden plus Eins. Umgekehrt läßt sich zeigen, daß $\mathfrak{M}^3$ durch die nichttrivialen irreduziblen Erzeugenden, die Henkelzahl, die Randsphärenzahl und den Orientierbarkeitscharakter eindeutig bestimmt ist (sofern im Falle einer orientierbaren 3-Mannigfaltigkeit M^3 orientierte Erzeugende angegeben werden), s. [6]. Diese Eindeutigkeitssätze beweisen wir anschließend an die Endlichkeitssätze.

HILFSSATZ 1. *M^3 sei eine beliebige (zusammenhängende, berandete oder unberandete) 3-Mannigfaltigkeit, Γ sei eine beliebige randtreue Zerlegung von M^3 in offene Raumelemente, Θ sei eine Normalzerlegung bezüglich Γ. $M_1^2, \ldots, M_r^2$ seien r zusammenhängende, miteinander punktfremde Normalflächen bezüglich Θ. Dann gilt:*

a) *Ist M'^3 eine z-Komponente von $M^3-(M_1^2+\cdots+M_r^2)$, die die 3-Elemente aus Θ nur in Normalprismen schneidet, so liegt einer der drei folgenden Fälle vor:*

I. *$\overline{M}'^3-M'^3$ besteht aus zwei Flächen $M_{i_1}^2$, $M_{i_2}^2$ (i_1, i_2 zwischen 1 und r), die zweiseitig in M^3 liegen und zueinander topologisch parallel sind.*

II. *$\overline{M}'^3-M'^3$ ist eine Fläche M_i^2 (i zwischen 1 und r), die zweiseitig in M^3 liegt. In diesem Falle gibt es eine einseitig in M'^3 liegende Normalfläche M^{*2}, so daß $\overline{M}'^3$ eine Umgebung von M^{*2} in M^3 ist und M'^3-M^{*2} die 3-Elemente aus Θ ebenfalls nur in Normalprismen schneidet.*

III. *$\overline{M}'^3 - M'^3$ besteht aus zwei Flächen $M^2_{i_1}$, $M^2_{i_2}$ (i_1, i_2 zwischen 1 und r), so daß $M^2_{i_1}$ einseitig in M^3 liegt und $M^2_{i_2}$ zweiseitig. In diesem Falle ist $\overline{M}'^3$ eine Umgebung von $M^2_{i_1}$ in M^3.*

b) *Die (nach [5] endliche und bestimmbare) Anzahl der P-Klassen bezüglich Θ werde mit g bezeichnet. Dann gilt:*

b 1. *Die Anzahl der z-Komponenten von $M^3 - (M^2_1 + \cdots + M^2_r)$ ist mindestens gleich $1 + \frac{r}{2} - g$; d.h. liegen in M^3 mehr als $2g$ miteinander punktfremde Normalflächen, so ist der Außenraum nicht zusammenhängend. Daraus folgt insbesondere, daß mehr als $2g$ einseitig in M^3 liegende Normalflächen nicht miteinander punktfremd sein können.*

b 2. *Ist $r > 6g$, so sind mindestens zwei der Flächen $M^2_1, \ldots, M^2_r$ topologisch parallel in M^3.*

Beweis. Die durch $M^2_1, \ldots, M^2_r$ bewirkte Unterteilung von Θ bezeichnen wir mit Ψ.

Beweis von a).

1. In M'^3 gibt es eine Normalfläche M^{*2}, die wir als *Mittelfläche von M'^3* bezeichnen, die jedes d-dimensionale Element E^d der durch Θ bewirkten Zerlegung $\Theta(M'^3)$ von M'^3 in genau einem $(d-1)$-dimensionalen Element F^{d-1} schneidet [$d = 1, 2, 3$; 0-dimensionale Elemente kommen in $\Theta(M'^3)$ nicht vor]. Man kann M^{*2} folgendermaßen erhalten: Zunächst zeichnet man in jedem 1-Element aus $\Theta(M'^3)$ einen Punkt aus und dann in jedem 2-Element E^2 aus $\Theta(M'^3)$ ein 1-Element, das von den beiden in $\dot{E}^2$ liegenden ausgezeichneten Punkten berandet wird [jedes 2-Element aus $\Theta(M'^3)$ enthält genau zwei 1-Elemente aus $\Theta(M'^3)$ im Rande]. Ist schließlich E^3 ein 3-Element (Normalprisma) aus $\Theta(M'^3)$, so bilden die in $\dot{E}^3$ liegenden 2- und 1-Elemente aus $\Theta(M'^3)$ einen offenen Kreisring und die darin ausgezeichneten 1- und 0-Elemente eine 1-Sphäre; dann zeichnen wir in E^3 ein 2-Element aus, das von dieser 1-Sphäre berandet wird. Die so ausgezeichneten Elemente bilden zusammen die Elemente der Zerlegung $\Theta(M^{*2})$.

2. Nun betrachten wir die Zuordnung α zwischen den Elementen aus $\Theta(M^{*2})$ und den Elementen aus $\Theta(\overline{M}'^3 - M'^3)$, die sich durch folgende Vorschrift ergibt: Liegt ein Element F^{d-1} aus $\Theta(M^{*2})$ in dem Element E^d aus $\Theta(M'^3)$, so werden ihm die beiden in $\dot{E}^d$ liegenden $(d-1)$-Elemente aus $\Theta(\overline{M}'^3 - M'^3)$ zugeordnet und umgekehrt. Diese Zuordnung α besitzt offenbar folgende Eigenschaften: A) Liegt F^e aus $\Theta(M^{*2})$ im Rande von F^d aus $\Theta(M^{*2})$ und sind F'^e und F''^e bzw. F'^d und F''^d die zugeordneten Elemente aus $\Theta(\overline{M}'^3 - M'^3)$, so liegt bei geeignet gewählter Bezeichnung F'^e in $\dot{F}'^d$ und F''^e in $\dot{F}''^d$ B) Einem Element F'^d aus $\Theta(\overline{M}'^3 - M'^3)$ sind entweder zwei Elemente oder ein Element aus $\Theta(M^{*2})$ zugeordnet, je nachdem die F'^d enthaltende z-Komponente von $\overline{M}'^3 - M'^3$ einseitig oder zweiseitig in M^3 liegt.

3. Da M^{*2} zusammenhängend ist, folgt aus der Eigenschaft A) von α, daß $\overline{M}'^3 - M'^3$ entweder ebenfalls zusammenhängend ist, oder aus genau zwei

z-Komponenten besteht, je nachdem M^{*2} einseitig oder zweiseitig in M^3 liegt. Andererseits enthält $\overline{M}'^3 - M'^3$ mindestens eine z-Komponente, die zweiseitig in M^3 liegt, da anderenfalls $\overline{M}'^3 = M^3$ wäre (es ist aber $\overline{M}'^3 \neq M^3$, da nicht alle 3-Elemente aus Ψ Normalprismen sind). Also sind nur drei Fälle möglich: I. $\overline{M}'^3 - M'^3$ besteht aus zwei zweiseitigen z-Komponenten $M^2_{i_1}$, $M^2_{i_2}$; II. $\overline{M}'^3 - M'^3 = M^2_i$ ist zusammenhängend und zweiseitig; III. $\overline{M}'^3 - M'^3$ besteht aus einer einseitigen z-Komponente $M^2_{i_1}$ und einer zweiseitigen $M^2_{i_2}$. Im ersten Falle umfaßt α eine isomorphe Abbildung von $\Theta(M^{*2})$ auf $\Theta(M^2_{i_1})$ und eine isomorphe Abbildung von $\Theta(M^{*2})$ auf $\Theta(M^2_{i_2})$; dabei läßt sich $M^2_{i_1}$ isotop in $M^2_{i_2}$ deformieren, so daß jedes $(d-1)$-Element aus $\Theta(M^2_{i_1})$ über dasjenige d-Element aus $\Theta(M'^3)$ verschoben wird, in dessen Rand es liegt. Im zweiten Falle liegt die Mittelfläche M^{*2} einseitig in M^3, und $\overline{M}'^3$ ist eine Umgebung von M^{*2} in M^3. Denn ist U^3 eine beliebige, im Verhältnis zu Ψ kleine Umgebung von M^{*2} in M^3, so ist $(\dot{U}^3 - \dot{M}^3)^-$ eine zweiseitige Normalfläche in M'^3, und die 3-Elemente aus $\Theta(M'^3 - U^3)$ sind sämtlich Normalprismen; also läßt sich $(\dot{U}^3 - \dot{M}^3)^-$ isotop über $M'^3 - U^3$ in $(\dot{M}'^3 - \dot{M}^3)^-$ verschieben, also ist $\overline{M}'^3$ ebenfalls eine Umgebung von M^{*2} in M^3. Im dritten Falle sind die 3-Elemente aus $\Theta(M'^3 + M^2_{i_1})$ ebenfalls sämtlich Normalprismen, und $M^2_{i_1}$ ist Mittelfläche von $M'^3 + M^2_{i_1}$; also ist (s. Fall II) $\overline{M}'^3$ eine Umgebung von $M^2_{i_1}$ in M^3. Damit ist a) bewiesen.

Beweis von b).

1. Die z-Komponenten von $M^3 - (M^2_1 + \cdots + M^2_r)$ bezeichnen wir mit $M^3_1, \ldots, M^3_s$ $(s \geqq 1)$. Aus der bereits bewiesenen Behauptung a) folgt, daß bei geeigneter Numerierung folgendes gilt:

A) $M^3_1, \ldots, M^3_t$ (t gegebenenfalls gleich Null) schneiden die 3-Elemente aus Θ nur in Normalprismen und enthalten je genau eine zweiseitig in M^3 liegende Normalfläche im Rande; dabei enthält jede dieser Komponenten eine einseitig in M^3 liegende Normalfläche.

B) $M^3_{t+1}, \ldots, M^3_{t+u}$ (u gegebenenfalls gleich Null) schneiden die 3-Elemente aus Θ nur in Normalprismen und enthalten je eine zweiseitige und je eine einseitige Normalfläche im Rande.

C) $M^3_{t+u+1}, \ldots, M^3_{t+u+v}$ (v gegebenenfalls gleich Null) schneiden die 3-Elemente aus Θ nur in Normalprismen und enthalten je zwei zweiseitige Normalflächen im Rande.

D) $M^3_{t+u+v+1}, \ldots, M^3_s$ $(s \geqq t+u+v+1)$ schneiden die 3-Elemente aus Θ nicht nur in Normalprismen.

2. Höchstens $2g$ der Flächen $M^2_1, \ldots, M^2_r$ liegen in $\dot{M}^3_{t+u+v+1} + \cdots + \dot{M}^3_s$.

BEWEIS. Die in $M^2_1 + \cdots + M^2_r$ liegenden P-Normalflächenstücke seien $P^2_1, \ldots, P^2_x$; unter diesen seien $P^2_1, \ldots, P^2_y$ diejenigen, die nicht zwischen zwei bezüglich Θ ähnlichen liegen, d.h. die nicht im Rande je zweier Normalprismen aus Ψ liegen. Offenbar ist $y \leqq 2g$.

Enthält eine Fläche M^2_i (i zwischen 1 und r) keines der Flächenstücke $P^2_1, \ldots, P^2_y$, so liegt auch jedes K- und jedes F-Normalflächenstück aus M^2_i

im Rande von je zwei Normalprismen aus Ψ. Wir betrachten nun alle diejenigen Normalprismen aus Ψ, die Flächenstücke aus M_i^2 im Rande enthalten; diese bilden (zusammen mit ihren nicht in $M_1^2+\cdots+M_r^2$ liegenden Randelementen) eine oder zwei z-Komponenten von $M^3-(M_1^2+\cdots+M_r^2)$, also eine oder zwei der z-Komponenten $M_1^3, \ldots, M_{t+u+v}^3$. Also kann M_i^2 nicht im Rande einer z-Komponente $M_{t+u+v+1}^3, \ldots, M_s^3$ liegen.

Also gibt es unter den Flächen $M_1^2, \ldots, M_r^2$ höchstens $y \leqq 2g$, die im Rande von Komponenten $M_{t+u+v+1}^3, \ldots, M_s^3$ liegen, w.z.b.w.

3. Unter den Flächen $M_1^2, \ldots, M_r^2$ werden die nicht in $M_{t+u+v+1}^3+\cdots+M_s^3$ liegenden mit $M_1'^2, \ldots, M_w'^2$ bezeichnet und nach Nr. 2 ist $w \geqq r-2g$. Da jede der Flächen $M_1'^2, \ldots, M_w'^2$ im Rande mindestens einer der z-Komponenten $M_1^3, \ldots, M_{t+u+v}^3$ liegt, und andererseits jede der z-Komponenten $M_1^3, \ldots, M_{t+u+v}^3$ [nach a)] höchstens zwei der Flächen $M_1'^2, \ldots, M_w'^2$ im Rande enthält, ist $t+u+v \geqq \frac{1}{2}w \geqq \frac{1}{2}r-g$. Also ist $s \geqq 1+\frac{1}{2}r-g$, womit b 1 bewiesen ist.

4. Da es nach der soeben bewiesenen Behauptung b 1 in M^3 höchstens $2g$ miteinander punktfremde Normalflächen geben kann, die einseitig in M^3 liegen, ist (s. Nr. 1) $t+u \leqq 2g$. Also ist nach Nr. 3 $v \geqq \frac{1}{2}r-3g$. Daraus folgt b 2; denn bei $r>6g$ ist $v>0$ und [nach a)] sind die beiden im Rande von M_{t+u+j}^3 $(j=1, \ldots, v)$ liegenden Normalflächen topologisch parallel. Damit ist Hilfssatz 1 bewiesen.

ENDLICHKEITSSATZ 1 (H. KNESER). *Ist M^3 eine beliebige (zusammenhängende, berandete oder unberandete) 3-Mannigfaltigkeit, so läßt sich eine Zahl g angeben, so daß folgendes gilt: Liegen in M^3 r beliebige, miteinander punktfremde 2-Sphären $S_1^2, \ldots, S_r^2$ und ist $r>6g$, so ist mindestens eine z-Komponente des Außenraums $M^3-(S_1^2+\cdots+S_r^2)$ eine gelochte 3-Sphäre, d.h. $S_1^2, \ldots, S_r^2$ sind geometrisch abhängig in M^3.*

Dabei kann die Zahl g folgendermaßen bestimmt werden: Es werden eine beliebige randtreue Zerlegung Γ von M^3 in offene Raumelemente und eine Normalzerlegung Θ bezüglich Γ gewählt. Dann ist g die (nach [5], Kap. II zu bestimmende) Anzahl der P-Klassen bezüglich Θ.

BEWEIS. Aus der Annahme, $S_1^2, \ldots, S_r^2$ seien geometrisch unabhängig und r sei größer als $6g$, folgt: Nach [5], Hauptsatz 1 läßt sich $S_1^2+\cdots+S_r^2$ durch ξ-Operationen in eine Normalfläche bezüglich Θ überführen, also existiert nach Satz 2 e) ein System von r miteinander punktfremden 2-Sphären $S_1'^2, \ldots, S_r'^2$, die Normalflächen bezüglich Θ sind und ebenfalls geometrisch unabhängig in M^3. Dieses steht im Widerspruch zu Hilfssatz 1 b 2, wonach zwei dieser 2-Sphären topologisch parallel sein müssen, also eine zweifach gelochte 3-Sphäre in M^3 beranden, also geometrisch abhängig sind. Also ist die Annahme falsch, w.z.b.w.

FOLGERUNG (H. KNESER). *Ist M^3 eine beliebige (zusammenhängende, berandete oder unberandete) 3-Mannigfaltigkeit, so existiert stets ein vollständiges Sphärensystem in M^3.*

BEWEIS. Gibt es in M^3 kein geometrisch unabhängiges Sphärensystem, so ist jedes beliebige (in $\mathring{M}^3$ liegende) Sphärensystem vollständig. — Existiert

in M^3 ein geometrisch unabhängiges Sphärensystem, so existiert nach Endlichkeitssatz 1 ein *maximales* geometrisch unabhängiges Sphärensystem in M^3, d.h. ein Sphärensystem M^2, so daß M^2+S^2 stets geometrisch abhängig ist, wenn S^2 eine beliebige 2-Sphäre in $\mathring{M}^3-M^2$ ist. Ist M^2 nicht einfach, so kann zu jeder 2-Sphäre aus M^2, die im Rande nur einer z-Komponente von M^3-M^2 liegt, eine topologisch parallele hinzugenommen werden. Hierdurch geht M^2 in ein vollständiges Sphärensystem über. Damit ist die Folgerung bewiesen.

ENDLICHKEITSSATZ 2. *Ist M^3 eine beliebige irreduzible (zusammenhängende, berandete oder unberandete) 3-Mannigfaltigkeit, so läßt sich eine Zahl g angeben, so daß folgendes gilt: Liegen in M^3 r beliebige, unberandete*[10]*, inkompressible, miteinander punktfremde 2-Mannigfaltigkeiten $M_1^2, \ldots, M_r^2$, so gilt:*

a) *Der Außenraum $M^3-(M_1^2+\cdots+M_r^2)$ zerfällt in mindestens $1+\frac{r}{2}-g$ z-Komponenten.*

b) *Ist $r>6g$, so sind mindestens zwei der 2-Mannigfaltigkeiten topologisch parallel in M^3.*

Dabei kann g ebenso bestimmt werden, wie in Endlichkeitssatz 1 angegeben.

BEWEIS. Nach [5], Hauptsatz 1 läßt sich $M_1^2+\cdots+M_r^2$ durch ξ-Operationen in eine Normalfläche überführen und daraus folgt nach Satz 2 a), daß es eine zu $M_1^2+\cdots+M_r^2$ isotope Normalfläche $M_1'^2+\cdots+M_r'^2$ gibt. Damit folgt der Endlichkeitssatz 2 aus Hilfssatz 1 b).

EINDEUTIGKEITSSATZ 1. *M^3 sei eine beliebige (zusammenhängende, berandete oder unberandete) 3-Mannigfaltigkeit; $S_1^2, \ldots, S_m^2$ und $S_1^{*2}, \ldots, S_{m^*}^{*2}$ seien zwei beliebige vollständige Sphärensysteme in M^3; $(M_1^3, \ldots, M_p^3)$ und $(M_1^{*3}, \ldots, M_{p^*}^{*3})$ seien zwei entsprechende Aufspaltungen von M^3 in irreduzible Erzeugende; dabei seien $M_1^3, \ldots, M_q^3$ nichttrivial und $M_{q+1}^3, \ldots, M_p^3$ seien trivial. Dann folgt:*

a) *Es ist $m^*+1-p^*=m+1-p$ gleich der Henkelzahl h von M^3.*

b) *Bei geeigneter Numerierung sind M_i^3 und M_i^{*3} kombinatorisch äquivalent für alle $i=1, \ldots, q$ und $M_{q+1}^{*3}, \ldots, M_{p^*}^{*3}$ sind trivial. Ist M^3 eine orientierte 3-Mannigfaltigkeit, so gilt diese Äquivalenzbeziehung auch für die entsprechend orientierten Erzeugenden.*

BEWEIS. 1. Zunächst läßt sich das Sphärensystem $S_1^2, \ldots, S_m^2$ durch kleine isotope Deformationen so abändern, daß es mit $S_1^{*2}, \ldots, S_{m^*}^{*2}$ höchstens Durchdringungslinien gemeinsam hat. Diese Durchdringungslinien lassen sich nun successive durch ϱ-Operationen beseitigen (vgl. den Beweis von Satz 1). Dadurch entsteht ein mit $S_1^{*2}, \ldots, S_{m^*}^{*2}$ punktfremdes Sphärensystem $S_1'^2, \ldots, S_{m'}'^2$, das nach Satz 3 ebenfalls vollständig ist und folgende Eigenschaften hat: Ist $(M_1'^3, \ldots, M_{p'}'^3)$ eine $S_1'^2+\cdots+S_{m'}'^2$ entsprechende Aufspaltung von M^3, so ist bei geeigneter Numerierung $M_i'^3$ zu M_i^3 kombinatorisch äquivalent für alle $i=1, \ldots, q$, $M_{q+1}'^3, \ldots, M_{p'}'^3$ sind trivial und $m'+1-p'$ ist gleich $m+1-p$.

[10] Nach geeigneter Verschärfung des Begriffes *inkompressibel* und entsprechender Verallgemeinerung des Hauptsatzes 1 aus [5] gilt dieser Endlichkeitssatz auch für berandete 2-Mannigfaltigkeiten, deren Ränder in $\dot{M}^3$ liegen.

2. Nun gilt allgemein: Ist $S_1''^2, \ldots, S_{m''}''^2$ ein vollständiges Sphärensystem, $(M_1''^3, \ldots, M_{p''}''^3)$ eine entsprechende Aufspaltung von M^3, und sind $S_{m''+1}''^2, \ldots, S_{m''+r}''^2$ miteinander und mit $S_1''^2, \ldots, S_{m''}''^2$ punktfremde 2-Sphären in $\mathring{M}^3$, so enthält eine $S_1''^2, \ldots, S_{m''+r}''^2$ entsprechende Aufspaltung von M^3 zu jeder 3-Mannigfaltigkeit $M_1''^3, \ldots, M_{p''}''^3$ eine äquivalente und außerdem weitere r triviale Erzeugende. Dieses ergibt sich leicht durch Induktion nach r.

3. Ist also $(M_1^{**3}, \ldots, M_{p\bullet}^{**3})$ eine $S_1'^2 + \cdots + S_{m'}'^2 + S_1^{*2} + \cdots + S_{m\bullet}^{*2}$ entsprechende Aufspaltung von M^3, so sind nach Nr. 2 bei geeigneter Numerierung einerseits $M_1^{**3}, \ldots, M_{p'}^{**3}$ äquivalent zu $M_1'^3, \ldots, M_{p'}'^3$ und andererseits $M_1^{**3}, \ldots, M_{p\bullet}^{**3}$ äquivalent zu $M_1^{*3}, \ldots, M_{p\bullet}^{*3}$. Daraus und aus Nr. 1 folgt, daß $M_1^{*3}, \ldots, M_q^{*3}$ kombinatorisch äquivalent zu $M_1^3, \ldots, M_q^3$ sind und $M_{q+1}^{*3}, \ldots, M_{p\bullet}^{*3}$ trivial, womit Behauptung b) bewiesen ist. Ferner ist $p^{**} = p' + m^* = p^* + m'$, also $m^* + 1 - p^* = m' + 1 - p' = m + 1 - p$, wie behauptet.

4. Es ist noch zu zeigen, daß $m+1-p$ die Henkelzahl h von M^3 ist:

4.1. An M^3 lassen sich $m+1-p$ Henkelreduktionen durchführen; denn unter den 2-Sphären $S_1^2, \ldots, S_m^2$ gibt es $m+1-p$ — bei geeigneter Numerierung sind diese $S_1^2, \ldots, S_{m+1-p}^2$ —, so daß $M^3 - (S_1^2 + \cdots + S_{m+1-p}^2)$ zusammenhängend ist.

Beweis. S_m^2 liegt (da das Sphärensystem einfach ist) im Rande zweier z-Komponenten des Außenraumes $M^3 - (S_1^2 + \cdots + S_m^2)$. Durch Fortlassen von S_m^2 entsteht ein Sphärensystem, dessen Außenraum aus $p-1$ z-Komponenten besteht. Ist $p-1>1$, so enthält dieses Sphärensystem (mindestens) eine weitere 2-Sphäre — bei geeigneter Numerierung ist diese S_{m-1}^2 —, die im Rande zweier z-Komponenten von $M^3 - (S_1^2 + \cdots + S_{m-1}^2)$ liegt. Dann lassen wir S_{m-1}^2 fort usw. und gelangen nach insgesamt $p-1$ Schritten zu dem Sphärensystem $S_1^2, \ldots, S_{m+1-p}^2$ mit den geforderten Eigenschaften, w.z.b.w.

4.2. An M^3 lassen sich nicht $m+2-p$ Henkelreduktionen durchführen.

Beweis. Aus der Annahme des Gegenteils folgt: Es gibt in M^3 $m+2-p$ miteinander punktfremde 2-Sphären $S_1'''^2, \ldots, S_{m+2-p}'''^2$, so daß $M^3 - (S_1'''^2 + \cdots + S_{m+2-p}'''^2)$ zusammenhängend ist. Ergänzen wir diese $m+2-p$ 2-Sphären durch weitere 2-Sphären zu einem vollständigen Sphärensystem $S_1'''^2, \ldots, S_{m'''}'''^2$ (das nach Endlichkeitssatz 1 stets existiert), so besteht $M^3 - (S_1'''^2 + \cdots + S_{m'''}'''^2)$ aus p''' z-Komponenten mit $p''' \leqq m''' - (m+2-p) + 1 = m''' + p - m - 1$, also ist $m''' + 1 - p''' > m + 1 - p$ im Widerspruch zu Nr. 3, wonach $m''' + 1 - p''' = m + 1 - p$ sein müßte, also ist die Annahme falsch, w.z.b.w.

Folgerung. *Ist M^3 eine beliebige (zusammenhängende, berandete oder unberandete) 3-Mannigfaltigkeit der Äquivalenzklasse $\mathfrak{M}^3$, so existieren endlich viele (gegebenenfalls Null) nichttriviale irreduzible Erzeugende von $\mathfrak{M}^3$ und diese sind (bis auf die Reihenfolge) eindeutig ausgezeichnet.* (Dieses ergibt sich unmittelbar aus der Existenz eines vollständigen Sphärensystems und aus Eindeutigkeitssatz 1.)

Eindeutigkeitssatz 2 (H. Kneser). *Sind M^3 und M'^3 beliebige (zusammenhängende, berandete oder unberandete) 3-Mannigfaltigkeiten, und sind*

$(M_1^3, \ldots, M_p^3)$ und $(M_1'^3, \ldots, M_{p'}'^3)$ Aufspaltungen von M^3 bzw. M'^3 in irreduzible Erzeugende, so sind M^3 und M'^3 dann (und nach Eindeutigkeitssatz 1 auch nur dann) kombinatorisch äquivalent, wenn folgendes gilt:

a) *M^3 und M'^3 besitzen gleichen Orientierbarkeitscharakter.*

b) *M^3 und M'^3 besitzen gleiche Henkelzahl und gleiche Randsphärenzahl.*

c) *Bei geeigneter Numerierung sind die Erzeugenden $M_1^3, \ldots, M_q^3$ nichttrivial und kombinatorisch äquivalent zu $M_1'^3, \ldots, M_q'^3$, und $M_{q+1}^3, \ldots, M_p^3, M_{q+1}'^3, \ldots, M_{p'}'^3$ sind trivial.*

d) *Sind M^3 und M'^3 orientierbar, so gilt bei geeigneter Orientierung von M^3 und M'^3 die unter c) geforderte kombinatorische Äquivalenz auch für die entsprechend orientierten Erzeugenden.*

BEWEIS. Gelten die Beziehungen a), b), c), d), sind h die Henkelzahl und l die Randsphärenzahl von M^3 und M'^3, und sind $R_1^2, \ldots, R_l^2$, bzw. $R_1'^2, \ldots, R_l'^2$ die Randsphären von M^3 und M'^3, so gilt:

1. In M^3 und in M'^3 existieren vollständige Sphärensysteme, die aus 2-Sphären $P_1^2, \ldots, P_h^2, Q_1^2, \ldots, Q_h^2, S_1^2, \ldots, S_q^2$ bzw. $P_1'^2, \ldots, P_h'^2, Q_1'^2, \ldots, Q_h'^2, S_1'^2, \ldots, S_q'^2$ bestehen und die M^3 bzw. M'^3 in z-Komponenten $H_1^3, \ldots, H_h^3, Z^3, Z_1^3, \ldots, Z_q^3$ bzw. $H_1'^3, \ldots, H_h'^3, Z'^3, Z_1'^3, \ldots, Z_q'^3$ zerlegen, so daß folgendes gilt:

I. Z^3 und Z'^3 sind $(2h+l+q)$-fach gelochte 3-Sphären. Es liegen also alle 2-Sphären $P_i^2, Q_i^2, S_j^2, R_k^2$ $(i=1, \ldots, h;\ j=1, \ldots, q;\ k=1, \ldots, l)$ bzw. $P_i'^2, Q_i'^2, S_j'^2, R_k'^2$ in $\dot{Z}^3$ bzw. $\dot{Z}'^3$.

II. In $\dot{Z}_j^3$ liegt die 2-Sphäre S_j^2 und keine weitere 2-Sphäre, in $\dot{Z}_j'^3$ liegt $S_j'^3$ und keine weitere 2-Sphäre, für alle $j=1, \ldots, q$. Dabei ist $\overline{Z}_j^3$ kombinatorisch äquivalent zu $\overline{Z}_j'^3$ (und ist kein 3-Element); sind M^3 und M'^3 orientierbar, so gilt diese Äquivalenz bei geeigneter Orientierung von M^3 und M'^3 auch für die entsprechend orientierten $\overline{Z}_j^3, \overline{Z}_j'^3$.

III. H_i^3 und $H_i'^3$ sind zweifach gelochte 3-Sphären mit $\dot{H}_i^3 = P_i^2 + Q_i^2$ und $\dot{H}_i'^3 = P_i'^2 + Q_i'^2$ für alle $i=1, \ldots, h$.

BEWEIS. In $\mathring{M}^3$ existieren h miteinander punktfremde 2-Sphären $P_1^2, \ldots, P_h^2$, so daß der Außenraum zusammenhängend ist. Zu jeder dieser 2-Sphären P_i^2 wählen wir in $\mathring{M}^3$ eine topologisch parallele 2-Sphäre Q_i^2. Die zwischen P_i^2 und Q_i^2 liegende (offene) zweifach gelochte 3-Sphäre bezeichnen wir mit H_i^3, die Punktmenge $M^3 - (\overline{H}_1^3 + \cdots + \overline{H}_h^3)$ mit $Z_{(1)}^3$.

Ist $Z_{(1)}^3$ eine gelochte 3-Sphäre, so bilden $P_1^2, Q_1^2, \ldots, P_h^2, Q_h^2$ ein vollständiges Sphärensystem in M^3 mit den geforderten Eigenschaften; $(Z^3 = Z_{(1)}^3)$. Nach Eindeutigkeitssatz 1 ist in diesem Falle $q=0$. — Ist $Z_{(1)}^3$ keine gelochte 3-Sphäre, so existiert in $\mathring{Z}_{(1)}^3$ eine 2-Sphäre S_1^2, die $Z_{(1)}^3$ in zwei z-Komponenten $Z_{(2)}^3$ und Z_1^3 zerlegt, so daß in $\dot{Z}_1^3$ keine 2-Sphären außer S_1^2 liegen und $\overline{Z}_1^3$ irreduzibel und kein 3-Element ist. Ist $Z_{(2)}^3$ keine gelochte 3-Sphäre, so existiert in $\mathring{Z}_{(2)}^3$ eine 2-Sphäre S_2^2 mit entsprechenden Eigenschaften, usw. Nach Endlichkeitssatz 1 existiert eine Zahl q^*, so daß $Z_{(q^*)}^3$ eine gelochte 3-Sphäre ist und

$P_1^2, Q_1^2, \ldots, P_k^2, Q_k^2, S_1^2, \ldots, S_{q^*}^2$ ein vollständiges Sphärensystem in M^3. Nach Eindeutigkeitssatz 1 ist daher $q^* = q$, und bei geeigneter Numerierung liefert die Ergänzung von Z_i^3 durch ein 3-Element eine zu M_i^3 äquivalente 3-Mannigfaltigkeit für alle $i = 1, \ldots, q$. Also besitzt das Sphärensystem $P_1^2, \ldots, S_q^2$ in M^3 die geforderten Eigenschaften ($Z^3 = Z_{(q)}^3$).

Genauso verfahren wir in M'^3 und erhalten ein vollständiges Sphärensystem $P_1'^2, \ldots, S_q'^2$. Die so erhaltenen Sphärensysteme besitzen die geforderten Eigenschaften, womit Nr. 1 bewiesen ist.

2. Wir behandeln nun den Fall, daß M^3 und M'^3 orientierbar sind: Dann existieren n.V. kohärente Orientierungen ω und ω' von M^3 und M'^3 und semilineare Abbildungen $\alpha_1, \ldots, \alpha_q$ von $\overline{Z}_1^3$ auf $\overline{Z}_1'^3, \ldots, \overline{Z}_q^3$ auf $\overline{Z}_q'^3$, die die ω bzw. ω' entsprechenden Orientierungen erhalten. Ferner seien $\beta_1, \ldots, \beta_k$ semilineare orientierungserhaltende Abbildungen von $\overline{H}_1^3$ auf $\overline{H}_1'^3, \ldots, \overline{H}_k^3$ auf $\overline{H}_k'^3$. Dann gibt es eine semilineare Abbildung α von $\overline{Z}^3$ auf $\overline{Z}'^3$, die für die 2-Sphären $S_1^2, \ldots, S_q^2, P_1^2, \ldots, Q_k^2$ mit $\alpha_1, \ldots, \alpha_q$ bzw. $\beta_1, \ldots, \beta_k$ übereinstimmt. Also bilden $\alpha, \alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_k$ zusammen eine semilineare Abbildung von M^3 auf M'^3, womit Eindeutigkeitssatz 2 für den orientierbaren Fall bewiesen ist.

3. Nun betrachten wir den Fall, daß M^3 und M'^3 nichtorientierbar sind und dabei (mindestens) eine der Erzeugenden nichtorientierbar ist. Es seien also $Z_1^3, \ldots, Z_r^3$ nichtorientierbar und $Z_{r+1}^3, \ldots, Z_q^3$ orientierbar. α_j sei eine semilineare Abbildung von $\overline{Z}_j^3$ auf $\overline{Z}_j'^3$ (für alle $j = 1, \ldots, q$).

Wir orientieren zunächst alle 3-Simplizes einer (hinreichend feinen) Simplizialzerlegung von M^3. Diese (nichtkohärente) Orientierung ω wählen wir so, daß folgendes gilt:

A) Z^3, alle H_i^3 und alle orientierbaren Z_j^3 sind kohärent orientiert.

B) In den nichtorientierbaren Z_j^3 ($j = 1, \ldots, r$) kehrt sich die Orientierung in 2-dimensionalen (abgeschlossenen) Punktmengen Y_j^2 um, wobei $Z_j^3 - Y_j^2$ zusammenhängend sind.

C) Außer in den Y_j^2 kehrt sich die Orientierung höchstens noch in gewissen 2-Sphären P_i^2 um.

Kehrt sich die Orientierung in einer der 2-Sphären P_i^2 um — z.B. in P_1^2 —, so beseitigen wir diese Umkehrung durch folgende Abänderung („Verlagerung") von Z_1^3 (s. Fig. 5): Wir wählen in $\mathring{Z}^3 + S_1^2 + \mathring{Z}_1^3$ ein 1-Element E^1, das einen Punkt aus S_1^2 mit einem Punkt aus P_1^2 verbindet, Y_1^2 in genau einem Punkte trifft (und zwar durchsetzt, wobei eine kleine Umgebung dieses Punktes in Y_1^2 ein 2-Element sein soll) und S_1^2 in einem weiteren Punkte trifft (durchsetzt). Der Y_1^2 mit P_1^2 verbindende Teil von E^1 sei E^{*1}. Nun sei U^3 eine kleine Umgebung von $\mathring{E}^1 \overline{Z}^3 + P_1^2$ in $\overline{Z}^3$, V^3 eine (im Verhältnis zu $Y_1^2 | S_1^2 | U^3$) kleine Umgebung von $\overline{E}^{*1} + P_1^2$ in $\overline{Z}^3 + \overline{Z}_1^3$ und W^3 eine (im Verhältnis zu $Y_1^2 | S_1^2 | V^3$) kleine Umgebung von $\overline{E}^1 + P_1^2$ in $\overline{Z}_1^3 + V^3$; X^3 sei die abgeschlossene Hülle der E^{*1} enthaltenden z-Komponente von $V^3 - Y_1^2$. Dann ersetzen wir $S_1^2 (\dot{U}^3 + \dot{W}^3)$ durch $[(\dot{U}^3 + \dot{W}^3) - (S_1^2 + P_1^2)]$ und $Y_1^2 \dot{X}^3$ durch $[\dot{X}^3 - (Y_1^2 + P_1^2)]$ und orientieren X^3 um. Dadurch wird die Orientierungsumkehrung bei P_1^2

aufgehoben, aus S_1^2 entsteht eine 2-Sphäre $S_1^{(1)2}$, aus Y_1^2 eine Punktmenge $Y_1^{(1)2}$. Entsprechend geht Z_1^3 in $Z_1^{(1)3} = [Z_1^3 + (S_1^2 \dot{U}^3)^\circ + \mathring{U}^3] - W^3$ über und Z^3 in $Z^{(1)3} = (Z^3 + Z_1^3) - \overline{Z}_1^{(1)3}$. Offenbar gibt es eine semilineare Abbildung $\gamma^{(1)}$ von $\overline{Z}_1^3$ auf $\overline{Z}_1^{(1)3}$, die Y_1^2 in $Y_1^{(1)2}$ überführt; (denn ist E^3 ein beliebiges, von P_1^2 berandetes 3-Element, dessen Inneres mit M^3 punktfremd ist, so läßt sich $\overline{Z}_1^3$ in $\overline{Z}^3 + \overline{Z}_1^3 + E^3$ isotop in $\overline{Z}_1^{(1)3}$ verschieben, so daß dabei Y_1^2 in $Y_1^{(1)2}$ übergeht). — So fahren wir fort, bis wir in M^3 zu einem Sphärensystem S_1^{*2}, $S_2^2, \ldots, S_q^2, P_1^2, Q_1^2, \ldots, P_h^2, Q_h^2$ mit z-Komponenten $Z^{*3}, Z_1^{*3}, Z_2^3, \ldots, Z_q^3, H_1^3, \ldots, H_h^3$ des Außenraumes und einer Orientierung ω^* gelangen, die sich nur in den Punktmengen Y_1^{*2} (in Z_1^{*3} liegend), $Y_2^2, \ldots, Y_r^2$ umkehrt. — Völlig analog gehen wir in M'^3 vor und verwenden die entsprechenden Bezeichnungen. Es gibt dann eine semilineare Abbildung α_1^* von $\overline{Z}_1^{*3}$ auf $\overline{Z}_1'^{*3}$.

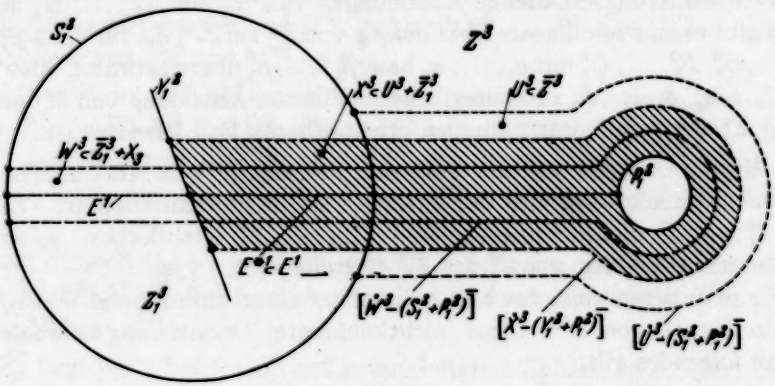

Fig. 5

Wir orientieren nun $\overline{Z}'^{*3} + H_1'^3 + \cdots + H_h'^3$ so, wie es ω'^* entspricht, und $Z_1'^{*3}, Z_2'^3, \ldots, Z_q'^3$ so, wie es der durch $\alpha_1^*, \alpha_2, \ldots, \alpha_q$ übertragenen Orientierung ω^* entspricht. Kehrt sich diese Orientierung in $S_1'^{*2}$ um, so beseitigen wir diese Umkehrung durch Umorientieren von $\overline{Z}'^{*3} + H_1'^3 + \cdots + H_h'^3$; kehrt sich dann die Orientierung in gewissen 2-Sphären $S_j'^2$ (j zwischen 2 und q) um, so beseitigen wir diese Umkehrungen successive nach dem im vorigen Absatz beschriebenen Verfahren durch Verlagerung von $Z_1'^{*3}$. So erhalten wir in M'^3 ein Sphärensystem $S_1'^{**2}, S_2'^2, \ldots, Q_h'^2$ und eine Orientierung ω'^{**}, die sich nur in Punktmengen $Y_1'^{**2}, Y_2'^{**2}, \ldots, Y_r'^{**2}$ umkehrt, die in den z-Komponenten $Z_1'^{**2}, Z_2'^3, \ldots, Z_r'^3$ des Außenraumes liegen. Dabei gibt es semilineare Abbildungen $\alpha_1^{**}, \alpha_2, \ldots, \alpha_q$ von $\overline{Z}_1^{*3}$ auf $\overline{Z}_1'^{**3}$, $\overline{Z}_2^3$ auf $\overline{Z}_2'^3, \ldots, \overline{Z}_q^3$ auf $\overline{Z}_q'^3$, durch die ω^* in ω'^{**} übergeht, Y_1^{*2} in $Y_1'^{**2}, \ldots, Y_r^2$ in $Y_r'^{**2}$. Nun gibt es semilineare Abbildungen β_i (für alle $i = 1, \ldots, h$) von $\overline{H}_i^3$ auf $\overline{H}_i'^3$, die ω^* in ω'^{**} übertragen, und eine semilineare Abbildung α von $\overline{Z}^{*3}$ auf $\overline{Z}'^{**3}$, die ebenfalls ω^* in ω'^{**} überführt und auf $S_1^{*2}, S_2^2, \ldots, S_q^2, P_1^2, Q_1^2, \ldots, P_h^2, Q_h^2$ mit $\alpha_1^{**}, \alpha_2, \ldots, \alpha_q, \beta_1, \ldots, \beta_h$ übereinstimmt. $\alpha, \alpha_1^{**}, \alpha_2, \ldots, \beta_h$ bilden zusammen eine semilineare Abbildung von M^3 auf M'^3. Damit ist Eindeutigkeitssatz 2 für den betrachteten Fall bewiesen.

4. Schließlich haben wir noch den Fall zu behandeln, daß M^3 und M'^3 nichtorientierbar, jedoch alle Erzeugenden orientierbar sind. Wir gehen ganz ähnlich vor, wie in Nr. 3 dargestellt: In M^3 wählen wir eine (nichtkohärente) Orientierung ω, so daß die Teilmannigfaltigkeiten $\overline{Z}^3 + \overline{Z}_1^3 + \cdots + \overline{Z}_q^3, \overline{H}_1^3, \ldots, \overline{H}_k^3$ jeweils kohärent orientiert sind und daß sich die Orientierung in den 2-Sphären $Q_1^2, \ldots, Q_k^2$ nicht umkehrt. Dann kehrt sich die Orientierung n.V. an mindestens einer der 2-Sphären P_i^2 um — bei geeigneter Numerierung an P_1^2. Kehrt sich die Orientierung noch an einer weiteren 2-Sphäre P_i^2 um — etwa an P_2^2 —, so beseitigen wir diese Orientierungsumkehrung durch eine Verlagerung von H_1^3 (s. Fig. 6): In $H_1^3 + Q_1^2 + \mathring{Z}^3$ wählen wir ein offenes 1-Element E^1, das einen Punkt aus P_1^2 mit einem Punkt aus P_2^2 verbindet und Q_1^2 in

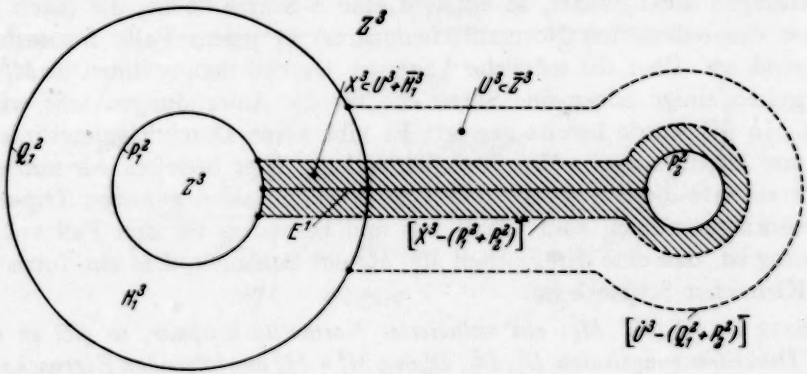

Fig. 6

genau einem Punkte trifft; U^3 sei eine kleine Umgebung von $\overline{E}^1 \overline{Z}^3 + P_2^2$ in $\overline{Z}^3$, X^3 eine kleine Umgebung von $\overline{E}^1 + P_2^2$ in $U^3 + \overline{H}_1^3$; dann ersetzen wir $\dot{U}^3 Q_1^2$ durch $[\dot{U}^3 - (Q_1^2 + P_2^2)]$, sowie $\dot{X}^3 P_1^2$ durch $[\dot{X}^3 - (P_1^2 + P_2^2)]$ und orientieren X^3 um. So fahren wir fort, bis wir ein Sphärensystem $P_1^{*2}, Q_1^{*2}, P_2^2, Q_2^2, \ldots, P_k^2, Q_k^2, S_1^2, \ldots, S_q^2$ (z-Komponenten des Außenraumes $Z^{*3}, H_1^{*3}, H_2^3, \ldots, Z_q^3$) und eine Orientierung ω^* erhalten, die sich nur in P_1^{*2} umkehrt. In M'^3 existiert ein entsprechendes Sphärensystem. — N.V. existieren semilineare Abbildungen α_i von $\overline{Z}_i^3$ auf $\overline{Z}_i'^3$. Nun orientieren wir $\overline{Z}'^{*3} + H_1'^{*3} + \cdots + H_k'^3$ so, wie es ω'^* entspricht, jedoch die $Z_i'^3$ so, wie es der durch α_i übertragenen Orientierung ω^* entspricht. Kehrt sich diese Orientierung in einer 2-Sphäre $S_i'^2$ um, so beseitigen wir diese Umkehrung durch Verlagerung von $H_1'^{*3}$, usw., bis wir in M'^3 ein Sphärensystem $P_1'^{**2}, Q_1'^{**2}, P_2'^2, Q_2'^2, \ldots, S_q'^2$ (z-Komponenten des Außenraumes $Z'^{**3}, H_1'^{**3}$, usw.) und eine Orientierung ω'^{**} erhalten, die sich nur in $P_1'^{**2}$ umkehrt und für alle $Z_j'^3$ mit der durch α_j übertragenen Orientierung ω^* übereinstimmt. Nun wählen wir semilineare Abbildungen $\beta_1, \ldots, \beta_k$ von $\overline{H}_1^{*3}$ auf $\overline{H}_1'^{**3}, \ldots, \overline{H}_k^3$ auf $\overline{H}_k'^3$, die ω^* in ω'^{**} überführen, und eine semilineare Abbildung α von $\overline{Z}^{*3}$ auf $\overline{Z}'^{**3}$, die ω^* in ω'^{**} überträgt und für die 2-Sphären $P_1^{*2}, \ldots, S_q^2$ mit $\beta_1, \ldots, \alpha_q$ übereinstimmt. Dann bilden $\alpha, \alpha_1, \ldots, \beta_k$ zusammen eine semilineare Abbildung von M^3 auf M'^3, womit Eindeutigkeitssatz 2 auch für den letzten Fall bewiesen ist.

§ 4. Reduktion der 3-dimensionalen Kreuzhauben und Henkel

In [5] wurde ausgeführt, daß jede Normalfläche umkehrbar eindeutig einer verträglichen nicht-negativ-ganzzahligen Lösung der P-Gleichungen entspricht; und zwar entspricht eine zusammenhängende Normalfläche M^2 entweder einer der endlich vielen nicht-negativ-ganzzahligen Fundamentallösungen der P-Gleichungen, oder es gibt ein reduziertes, M^2 erzeugendes Normalflächenpaar (M_1^2, M_2^2). Die Normalflächen M_1^2 und M_2^2 durchdringen einander und bilden zusammen eine S-Normalfläche. Längs jeder Durchdringungslinie gibt es eine eindeutig bestimmte Umschaltung, die als regulär bezeichnet wird, so daß durch Ausführung aller dieser Umschaltungen eine zu M^2 ähnliche Normalfläche entsteht. Wird eine beliebige Teilmenge der regulären Umschaltungen durchgeführt, so entsteht eine S-Normalfläche, die (nach Definition des reduzierten Normalflächenpaares) in jedem Falle S-zusammenhängend ist. Über die mögliche Lage der Durchdringungslinien in M_1^2 und M_2^2 gelten einige allgemeine Sätze, die für die Anwendungen sehr wichtig sind. In [5] wurde bereits gezeigt: Es gibt keine Durchdringungslinie, die M_1^2 und M_2^2 in je zwei z-Komponenten zerlegt. Hier beweisen wir nun einen weiteren Satz dieser Art, der das Nicht-Vorhandensein gewissen Tripel von Durchdringungslinien zum Inhalt hat und besonders für den Fall von Bedeutung ist, daß eine der Flächen M_1^2, M_2^2 ein Schlauch, d.h. ein Torus oder ein Kleinscher Schlauch ist.

SATZ 4. *Ist (M_1^2, M_2^2) ein reduziertes Normalflächenpaar, so gibt es keine drei Durchdringungslinien D_1^1, D_2^1, D_3^1 von $M_1^2 + M_2^2$ mit folgenden Eigenschaften:*

a) *Jede der Durchdringungslinien D_1^1, D_2^1, D_3^1 liegt im Rande zweier z-Komponenten von $M_1^2 - (D_1^1 + D_2^1 + D_3^1)$ und im Rande zweier z-Komponenten von $M_2^2 - (D_1^1 + D_2^1 + D_3^1)$.*

b) *Jede der z-Komponenten von $M_1^2 - (D_1^1 + D_2^1 + D_3^1)$ und von $M_2^2 - (D_1^1 + D_2^1 + D_3^1)$ enthält höchstens zwei Durchdringungslinien im Rande.*

BEWEIS. Aus der Annahme, es gäbe drei Durchdringungslinien D_1^1, D_2^1, D_3^1 mit den genannten Eigenschaften, folgt:

$M_i^2 - (D_1^1 + D_2^1 + D_3^1)$ besteht aus drei oder aus vier z-Komponenten (für $i = 1, 2$). Denn wegen der Eigenschaft a) sind mindestens zwei z-Komponenten vorhanden, und die Annahme, es seien genau zwei z-Komponenten vorhanden, widerspricht der Eigenschaft b); also gibt es mindestens drei z-Komponenten; andererseits gibt es höchstens vier z-Komponenten, da M_i^2 n.V. zusammenhängend ist. Sind vier z-Komponenten vorhanden, so wird M_i^2 durch jede einzelne der Durchdringungslinien bereits in zwei z-Komponenten zerlegt; da es aber (nach [5], Kap. V.1) keine Durchdringungslinie gibt, die sowohl M_1^2 als auch M_2^2 in je zwei z-Komponenten zerlegt, treten also nicht für $i = 1$ und $i = 2$ je vier z-Komponenten auf; d.h. bei geeigneter Numerierung besteht $M_1^2 - (D_1^1 + D_2^1 + D_3^1)$ aus genau drei z-Komponenten. Diese z-Komponenten können wir mit K_1^2, K_2^2 und K_3^2 bezeichnen, so daß D_1^1 in $\dot{K}_1^2$ und $\dot{K}_3^2$ liegt, D_2^1 in $\dot{K}_1^2$ und $\dot{K}_2^2$, D_3^1 in $\dot{K}_2^2$ und $\dot{K}_3^2$, s. Fig. 7. $M_2^2 - (D_1^1 + D_2^1 + D_3^1)$ besteht aus drei oder vier z-Komponenten; in jedem Falle gibt es (mindestens) zwei

unter ihnen — sie mögen mit Z_1^2 und Z_2^2 bezeichnet werden —, die von je zwei Durchdringungslinien berandet werden. Genau eine der Durchdringungslinien liegt in $\dot{Z}_1^2$ und in $\dot{Z}_2^2$; durch geeignete Numerierung kann erreicht werden, daß diese D_2^1 ist, daß ferner die von D_2^1 verschiedene, in $\dot{Z}_1^2$ liegende Durchdringungslinie D_1^1 ist und entsprechend die in $\dot{Z}_2^2$ liegende D_3^1.

Nun seien U_1^3, U_2^3 und U_3^3 kleine Umgebungen von D_1^1, D_2^1 bzw. D_3^1 in M^3. Die Durchschnitte einer solchen Umgebung mit M_1^2 und M_2^2 können keine Möbiusschen Bänder sein (da die Durchdringungslinien n.V. im Rande je zweier z-Komponenten liegen). Also besteht $U_j^3 - (M_1^2 + M_2^2 + M^3)$ ($j = 1, 2$ und 3) aus vier z-Komponenten $X_j^2, Y_j^2, V_j^2, W_j^2$, so daß folgendes gilt: (s. Fig. 7) $\overline{V}^2$ ist mit $\overline{W}_j^2$ punktfremd, X_j^2 mit $\overline{Y}_j^2$; V_1^2 verbindet K_3^2 und Z_1^2 mit-

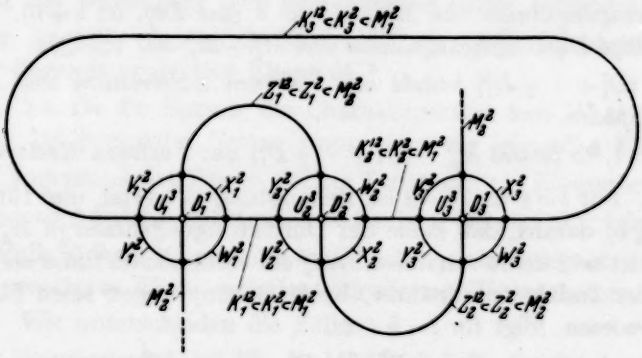

Fig. 7

einander, V_2^2 verbindet K_1^2 mit Z_2^2, V_3^2 K_2^2 mit $M_2^2 - Z_2^2$, X_1^2 K_1^2 mit Z_1^2, X_2^2 K_2^2 mit Z_2^2, X_3^2 K_3^2 mit $M_2^2 - Z_2^2$. Die Ersetzung von $(M_1^2 + M_2^2)\, U_j^3$ ($j = 1, 2, 3$) durch $\overline{V}_j^2 + \overline{W}_j^2$ ist eine Umschaltung, die wir mit σ_j bezeichnen; die entgegengesetzte Umschaltung bezeichnen wir mit $\bar{\sigma}_j$. Ferner bezeichnen wir die außerhalb von U_1^2, U_2^2 und U_3^2 liegenden Teile von K_j^2 und Z_i^2 mit $K_j'^2$ bzw. $Z_i'^2$.

Jeweils genau eine der Umschaltungen σ_j, $\bar{\sigma}_j$ ist regulär bezüglich $M_1^2 + M_2^2$ (s. [5], Hauptsatz 4). Hierbei gibt es acht mögliche Fälle:

a.a.a. σ_1, σ_2, σ_3 sind regulär,
a.a.b. σ_1, σ_2, $\bar{\sigma}_3$ sind regulär,
. . . . . .
b.b.b. $\bar{\sigma}_1$, $\bar{\sigma}_2$, $\bar{\sigma}_3$ sind regulär.

In allen acht Fällen folgt nun ein Widerspruch zu der Voraussetzung, daß (M_1^2, M_2^2) ein reduziertes Normalflächenpaar sei (vgl. Fig. 7):

a.a.a. und a.a.b: $K_3'^2 + \overline{V}_1'^2 + Z_1'^2 + \overline{W}_2^2 + K_2'^2 + M_1^2 U_3^3$ ist eine Sz-Komponente von $\sigma_1 \sigma_2 (M_1^2 + M_2^2)$, also ist $\sigma_1 \sigma_2 (M_1^2 + M_2^2)$ nicht S-zusammenhängend; dieses bedeutet in den Fällen a.a.a und a.a.b einen Widerspruch zur Voraussetzung.

a.b.a: $K_3'^2 + \overline{V}_1'^2 + Z_1'^2 + M_2^2 U_2^3 + Z_2'^2 + \overline{W}_3^2$ ist eine Sz-Komponente von $\sigma_1 \sigma_3 (M_1^2 + M_2^2)$.

a.b.b und b.b.b: $K_2'^2 + \bar{X}_2^2 + Z_2'^2 + \bar{Y}_3^2$ ist Sz-Komponente von $\bar{\sigma}_2\bar{\sigma}_3(M_1^2 + M_2^2)$.

b.a.a: $K_3'^2 + M_1^3 U_1^3 + K_1'^2 + \bar{V}_2^2 + Z_2'^2 + \bar{W}_3^2$ ist Sz-Komponente von $\sigma_2\sigma_3(M_1^2 + M_2^2)$.

b.a.b: $K_3'^2 + \bar{Y}_1^2 + \bar{X}_3^2 + [M_2^2 - (Z_1^2 + Z_2^2 + U_1^3 + U_3^3)]$ ist Sz-Komponente von $\bar{\sigma}_1\bar{\sigma}_3(M_1^2 + M_2^2)$.

b.b.a: $Z_1'^2 + \bar{X}_1^2 + K_1'^2 + \bar{Y}_2^2$ ist Sz-Komponente von $\bar{\sigma}_1\bar{\sigma}_2(M_1^2 + M_2^2)$.

Also ist die Annahme, es gäbe drei Durchdringungslinien mit den genannten Eigenschaften, falsch und Satz 4 damit bewiesen.

FOLGERUNG. *(M_1^2, M_2^2) sei ein reduziertes Normalflächenpaar; dabei sei M_1^2 eine 2-Sphäre und M_2^2 ein Torus oder ein Kleinscher Schlauch. Die Anzahl der Durchdringungslinien von $M_1^2 + M_2^2$ sei s (per Def. ist $s \neq 0$). $D_1^1, \ldots, D_t^1$ seien t beliebige Durchdringungslinien von $M_1^2 + M_2^2$ mit $1 \leqq t \leqq s$.* Dann folgt:

a) *$M_1^2 - (D_1^1 + \cdots + D_t^1)$ besteht aus t offenen 2-Elementen und einer t-fach gelochten 2-Sphäre.*

b) *Ist $t > 1$, so besteht $M_2^2 - (D_1^1 + \cdots + D_t^1)$ aus t offenen Kreisringen.*

BEWEIS. Für $t = 1$ und $t = 2$ ist Behauptung a) trivial, und für $t = 2$ folgt Behauptung b) daraus, daß keine der Durchdringungslinien in M_2^2 ein 2-Element berandet und keine von ihnen in M_2^2 ein Möbiussches Band als Umgebung hat. Aus der Induktionsannahme, beide Behauptungen seien für $t = u$ mit $1 < u < s$ bewiesen, folgt für $t = u + 1$:

Die u Durchdringungslinien $D_1^1, \ldots, D_u^1$ zerlegen M_2^2 in u offene Kreisringe; da D_{u+1}^1 in M_2^2 kein 2-Element berandet, zerlegt sie einen der Kreisringe in zwei Kreisringe, womit Behauptung b) auch für $t = u + 1$ bewiesen ist.

$D_1^1, \ldots, D_u^1$ zerlegen M_1^2 in u 2-Elemente $E_1^2, \ldots, E_u^2$ und eine u-fach gelochte 2-Sphäre T^2. Dann liegt D_{u+1}^1 nicht in einem der 2-Elemente E_i^2, da anderenfalls die drei Durchdringungslinien $D_{u+1}^1, \dot{E}_i^2, \dot{E}_j^2$ ($j \neq i$ beliebig) so in M_1^2 und M_2^2 lägen, wie es nach Satz 4 unmöglich ist. D_{u+1}^1 liegt also in T^2 und zerlegt T^2 in zwei z-Komponenten T_1^2 und T_2^2. Dabei ist eine dieser z-Komponenten ein offenes 2-Element; denn läge in T_1^2 eine Durchdringungslinie D_i^1 und in T_2^2 eine Durchdringungslinie D_j^1 mit i, j zwischen 1 und u, so lägen D_i^1, D_j^1, D_{u+1}^1 so in M_1^2 und M_2^2, wie es nach Satz 4 unmöglich ist. Damit ist auch Behauptung a) für $t = u + 1$ bewiesen.

Nach diesen Ergänzungen zur allgemeinen Theorie der Normalflächen können wir nun die Bestimmungsverfahren für die Kreuzhauben und Henkel von M^3 herleiten.

SATZ 5. *M^3 sei eine beliebige (zusammenhängende, berandete oder unberandete) 3-Mannigfaltigkeit. Θ sei eine Normalzerlegung bezüglich einer beliebigen randtreuen Zerlegung Γ von M^3 in offene Raumelemente. Dann gilt: Gibt es in M^3 eine (singularitätenfreie) einseitig in M^3 liegende projektive Ebene, so sind die Repräsentanten mindestens einer der Fundamentalklassen bezüglich Θ ebenfalls einseitig in M^3 liegende projektive Ebenen.*

BEWEIS. 1. Gibt es in M^3 eine einseitige projektive Ebene, so läßt sich diese (nach [5], Hauptsatz 1) durch ξ-Operationen in eine Normalfläche bezüglich Θ überführen. Daraus folgt nach Satz 2 b), daß es eine Normalfläche M^2 gibt, die ebenfalls eine einseitige projektive Ebene ist. Ihre F-Zahl sei f.

2. Liegt M^2 bereits in einer der Fundamentalklassen, so ist nichts mehr zu beweisen. Liegt M^2 in keiner der Fundamentalklassen, so gibt es nach [5], Satz 7 ein reduziertes, M^2 erzeugendes Normalflächenpaar (M_1^2, M_2^2). Die Durchdringungslinien von $M_1^2 + M_2^2$ seien $D_1^1, \ldots, D_s^1$; $U_1^3, \ldots, U_s^3$ seien im Verhältnis zu $\Theta | M_1^2 + M_2^2$ kleine Umgebungen von $D_1^1, \ldots, D_s^1$ in M^3. Die (nach [5], Hauptsatz 4) eindeutig bestimmten, bezüglich $M_1^2 + M_2^2$ regulären Umschaltungen von $(M_1^2 + M_2^2)\, U_i^3$ über $\dot{U}_i^3$ $(i = 1, \ldots, s)$ bezeichnen wir mit σ_i. Nun besteht der wesentliche Teil des Beweises darin, folgendes zu zeigen: Es gibt eine irregulär (bezüglich $\dot{U}_1^3, \ldots, \dot{U}_s^3$) aus $M_1^2 + M_2^2$ abgeleitete, einseitig in M^3 liegende projektive Ebene M'^2.

BEWEIS. 2.1. Da die Summe der Charakteristiken von M_1^2 und M_2^2 gleich -1 ist, liegt bei geeigneter Numerierung einer der folgenden Fälle vor:

a) M_1^2 ist eine projektive Ebene, M_2^2 ein Torus oder ein Kleinscher Schlauch.

b) M_1^2 ist eine 2-Sphäre mit drei Kreuzhauben, M_2^2 eine 2-Sphäre.

2.2. Im Falle a) liegt M_1^2 einseitig in M^3, ist also eine irregulär aus $M_1^2 + M_2^2$ abgeleitete projektive Ebene mit den geforderten Eigenschaften.

BEWEIS. Wir unterscheiden die Fälle:

a 1. Alle Durchschnitte $M_1^2\, U_1^3, \ldots, M_1^2\, U_s^3$ sind Kreisringe.

a 2. Es gibt ein j zwischen 1 und s, so daß $M_1^2\, U_j^3$ ein Möbiussches Band ist.

Im Falle a 1 enthält $M_1^2 - (U_1^3 + \cdots + U_s^3)$ eine z-Komponente P^2, die eine gelochte projektive Ebene ist. P^2 liegt in der zu M^2 ähnlichen, also einseitig in M^3 liegenden, projektiven Ebene $\sigma_s \ldots \sigma_1 (M_1^2 + M_2^2)$. Da die z-Komponenten von $\sigma_s \ldots \sigma_1 (M_1^2 + M_2^2) - P^2$ 2-Elemente sind, also zweiseitig in M^3 liegen, folgt daraus, daß P^2 einseitig in M^3 liegt. Also liegt auch M_1^2 einseitig in M^3, w.z.b.w.

Im Falle a 2 ist $M_2^2\, U_j^3$ ebenfalls ein Möbiussches Band (s. [5]), also liegt $M_1^2\, U_j^3$ einseitig in M^3, also auch M_1^2 selbst, w.z.b.w.

2.3. Im Falle b) sind alle Durchschnitte $M_1^2\, U_1^3, \ldots, M_1^2\, U_s^3$ Kreisringe, und keine der Durchdringungslinien zerlegt M_1^2 in zwei z-Komponenten. Also besteht $M_1^2 - (U_1^3 + \cdots + U_s^3)$ aus einer zweifach gelochten projektiven Ebene P^2 und $s-1$ Kreisringen. (Diese Folgerung ergibt sich leicht durch Betrachtung der Charakteristiken: Alle z-Komponenten sind mindestens zweifach gelocht, haben also nicht-negative Charakteristiken; also kommt genau eine z-Komponente mit der Charakteristik $+1$ vor und außer dieser nur Kreisringe.) Dabei liegt P^2 einseitig in M^3, da P^2 in der zu M^2 ähnlichen, einseitigen projektiven Ebene $\sigma_s \ldots \sigma_1 (M_1^2 + M_2^2)$ liegt. Nun unterscheiden wir die Fälle $s = 1$ und $s > 1$.

Im Falle $s = 1$ gilt: $M_2^2 - U_1^3$ besteht aus zwei 2-Elementen E_1^2, E_2^2. $\dot{U}_1^3 - (M_1^2 + M_2^2)$ besteht aus vier Kreisringen V^2, W^2, X^2, Y^2, so daß $\overline{W}^2$ mit

$\overline{V}^2$ punktfremd ist und $\overline{Y}^2$ mit $\overline{X}^2$. Bei geeigneter Bezeichnung sind $\overline{V}^2$ und $\overline{X}^2$ mit E_2^2 punktfremd, entsprechend $\overline{W}^2$ und $\overline{Y}^2$ mit E_1^2 (s. Fig. 8), und die Umschaltung σ_1 besteht in der Ersetzung von $(M_1^2+M_2^2)\,U_1^3$ durch $\overline{V}^2+\overline{W}^2$. Nun ist auch $\bar{\sigma}_1(M_1^2+M_2^2)=P^2+\overline{X}^2+E_1^2+\overline{Y}^2+E_2^2$ eine einseitige projektive Ebene; sie ist irregulär aus $M_1^2+M_2^2$ abgeleitet, womit die Behauptung für $s=1$ bewiesen ist.

Im Falle $s>1$ gilt: Die beiden Randlinien L_1^1 und L_2^1 von P^2 liegen im Rande zweier Umgebungen $U_{j_1}^3$ bzw. $U_{j_2}^3$ $(j_1, j_2=1, \ldots, s; j_1 \neq j_2)$. Dabei besteht $M_2^2-(U_{j_1}^3+U_{j_2}^3)$ aus zwei 2-Elementen E_1^2, E_2^2 und einem Kreisring K^2 (s. Fig. 9). $\dot{U}_{j_l}^3-(M_1^2+M_2^2)$ $(l=1, 2)$ besteht aus vier Kreisringen, deren einer

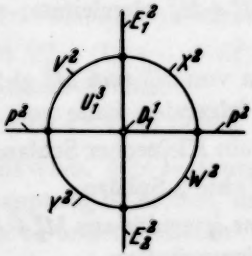

Fig. 8

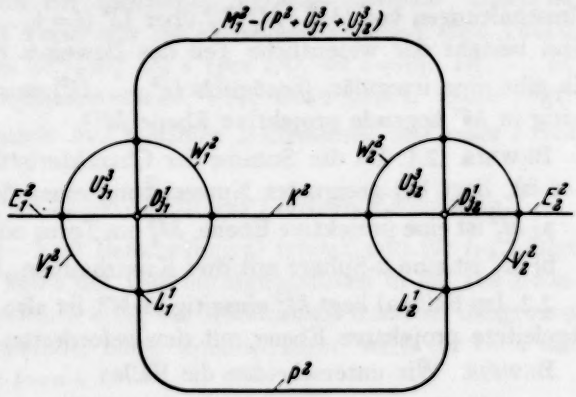

Fig. 9

— er werde mit V_l^2 bezeichnet — von L_l^1 und $\dot{E}_l^2$ berandet wird; der Kreisring, dessen abgeschlossene Hülle mit $\overline{V}_l^2$ punktfremd ist, wird mit $\overline{W}_l^2$ bezeichnet. Wir bezeichnen diejenige Umschaltung von $(M_1^2+M_2^2)\,U_l^3$ über $\dot{U}_l^3$, bei der $(M_1^2+M_2^2)\,U_l^3$ durch $\overline{V}_l^2+\overline{W}_l^2$ ersetzt wird, mit σ_l^*. Nun ist $P^2+\overline{V}_1^2+\overline{V}_2^2+E_1^2+E_2^2$ eine Sz-Komponente von $\sigma_2^*\sigma_1^*(M_1^2+M_2^2)$ und eine einseitige projektive Ebene. Diese projektive Ebene ist von $\sigma_s \ldots \sigma_1(M_1^2+M_2^2)$ verschieden, (da sie z. B. K^2 nicht enthält); also ist sie irregulär aus $M_1^2+M_2^2$ abgeleitet, womit die Behauptung für $s>1$ bewiesen ist.

Mit 2.1, 2.2 und 2.3 ist Nr. 2 bewiesen.

3. M'^2 ist eine Halbnormalfläche, und es ist entweder die F-Zahl f' von M'^2 kleiner als f, oder es ist $f'=f$ und M'^2 enthält eine Falte (s. [5], Beweis von Satz 8, Nr. 1). Daraus folgt nach [5], Hauptsatz 1 und nach Satz 2 b), daß es eine Normalfläche M^{*2} gibt, die ebenfalls eine einseitige projektive Ebene ist, und deren F-Zahl f^* kleiner als f ist.

4. Liegt M^{*2} in keiner der Fundamentalklassen, so läßt sich wieder, wie unter Nr. 2 und 3 ausgeführt, auf eine einseitige projektive Ebene M^{**2} schließen, die eine Normalfläche mit F-Zahl $f^{**}<f^*$ ist. So fortfahrend, kommen wir nach weniger als f Schritten zu einer einseitigen projektiven Ebene, die in einer der Fundamentalklassen liegt. Damit ist Satz 5 bewiesen.

SATZ 6. *M^3 sei eine kreuzhaubenfreie*[11]*, im übrigen beliebige (zusammenhängende, berandete oder unberandete) 3-Mannigfaltigkeit. Θ sei eine Normalzerlegung bezüglich einer beliebigen randtreuen Zerlegung Γ von M^3 in offene Raumelemente. Dann gilt: Gibt es in M^3 eine nichtzerlegende 2-Sphäre, so sind die Repräsentanten mindestens einer der Fundamentalklassen bezüglich Θ ebenfalls nichtzerlegende 2-Sphären in M^3.*

BEWEIS. 1. Gibt es in M^3 eine nichtzerlegende 2-Sphäre, so folgt nach [5], Hauptsatz 1 und nach Satz 2 c), daß es eine nichtzerlegende 2-Sphäre S^2 gibt, die eine Normalfläche bezüglich Θ ist. Die F-Zahl von S^2 sei f.

2. Liegt S^2 in keiner der Fundamentalklassen, so gibt es nach [5], Satz 7 ein reduziertes, S^2 erzeugendes Normalflächenpaar (M_1^2, M_2^2). Die Durchdringungslinien von $M_1^2+M_2^2$ seien $D_1^1, \ldots, D_s^1$; $U_1^3, \ldots, U_s^3$ seien kleine Umgebungen von $D_1^1, \ldots, D_s^1$ in M^3; $\sigma_1, \ldots, \sigma_s$ seien die entsprechenden, bezüglich $M_1^2+M_2^2$ regulären Umschaltungen. Dann gibt es eine irregulär aus $M_1^2+M_2^2$ abgeleitete, nichtzerlegende 2-Sphäre S'^2.

BEWEIS. 2.1. Die Summe der Charakteristiken von M_1^2 und von M_2^2 ist -2; also ist bei geeigneter Numerierung M_1^2 eine 2-Sphäre, M_2^2 ein Torus oder ein Kleinscher Schlauch. Denn daß M_1^2 und M_2^2 projektive Ebenen sind und damit n.V. zweiseitig in M^3 liegen, ist unmöglich, da dann jede Durchdringungslinie D_i^1 sowohl M_1^2 als auch M_2^2 in je zwei z-Komponenten zerlegen würde, (da $U_i^3 M_j^2$ kein zweiseitiges Möbiussches Band sein kann, also ein Kreisring ist).

2.2. Liegt M_1^2 nichtzerlegend in M^3, so besitzt M_1^2 die für S'^2 geforderten Eigenschaften. Wir haben daher nur noch den Fall zu behandeln, daß M_1^2 zerlegend in M^3 liegt. In diesem Falle gilt weiter:

2.3. Nach der Folgerung aus Satz 4 besteht $M_2^2-(U_1^3+\cdots+U_s^3)$ aus s offenen Kreisringen $K_1^2, \ldots, K_s^2$; da $M^3-M_1^2$ aus zwei z-Komponenten besteht, ist s eine gerade Zahl. Bei geeigneter Numerierung liegt eine Randlinie von K_i^2 in $\dot{U}_{i-1}^3$, die andere in $\dot{U}_i^3$ (Indices modulo s). Ferner besteht $M_1^2-(U_1^3+\cdots+U_s^3)$ aus s offenen 2-Elementen $E_1^2, \ldots, E_s^2$ und einer s-fach gelochten 2-Sphäre A^2; bei geeigneter Numerierung liegt $\dot{E}_i^2$ in $\dot{U}_i^3$ (s. Fig. 10). $\dot{U}_i^3-(M_1^2+M_2^2)$ besteht aus vier offenen Kreisringen $V_i^2, W_i^2, X_i^2, Y_i^2$; dabei kann die Bezeichnung so gewählt werden, daß $\overline{W}_i^2$ mit $\overline{V}_i^2$ punktfremd ist und entsprechend $\overline{Y}_i^2$ mit $\overline{X}_i^2$, daß $\dot{E}_i^2$ in $\dot{W}_i^2$ und in $\dot{X}_i^2$ liegt und daß die Umschaltung σ_i in der Ersetzung von $(M_1^2+M_2^2)U_i^3$ durch $\overline{V}_i^2+\overline{W}_i^2$ besteht. Dann liegt genau eine Randlinie von V_i^2 in $\dot{A}^2$. Durch geeignete Wahl des Numerierungssinnes i kann erreicht werden, daß die nicht in $\dot{A}^2$ liegende Randlinie von V_1^2 in $\dot{K}_1^2$ liegt und entsprechend die von $\dot{E}_1^2$ verschiedene Randlinie von W_1^2 in $\dot{K}_2^2$. Dann folgt durch Induktion, daß für alle i die nicht in $\dot{A}^2$ liegende Randlinie von V_i^2 in $\dot{K}_i^2$ liegt und die von $\dot{E}_i^2$ verschiedene Randlinie von W_i^2

[11]) Die Voraussetzung der Kreuzhaubenfreiheit ist, wie H. SCHUBERT bemerkt hat, leicht zu vermeiden.

in $\dot{K}^2_{i+1}$. Denn anderenfalls wäre $E^2_{j-1}+\overline{W}^2_{j-1}+K^2_j+\overline{W}^2_j+E^2_j$ eine z-Komponente von $\sigma_s \dots \sigma_1(M^2_1+M^2_2)$, wenn j die kleinste Zahl ist, für die nicht V^2_j, sondern W^2_j eine Randlinie aus $\dot{K}^2_j$ enthält; also wäre $\sigma_s \dots \sigma_1(M^2_1+M^2_2)$ nicht zusammenhängend.

2.4. Es gibt einen Index j zwischen 1 und s, so daß die 2-Sphäre $S'^2 = E^2_{j-1}+\overline{W}^2_{j-1}+K^2_j+\overline{X}^2_j+E^2_j$ nichtzerlegend in M^3 liegt (vgl. Fig. 10).

BEWEIS. Aus der Annahme des Gegenteils folgt: Für alle $i=1, \dots, s$ besteht $M^3-(E^2_{i-1}+\overline{W}^2_{i-1}+K^2_i+\overline{X}^2_i+E^2_i)$ aus zwei z-Komponenten, deren eine — sie werde mit T^3_i bezeichnet — die gelochte 2-Sphäre A^2 nicht enthält.

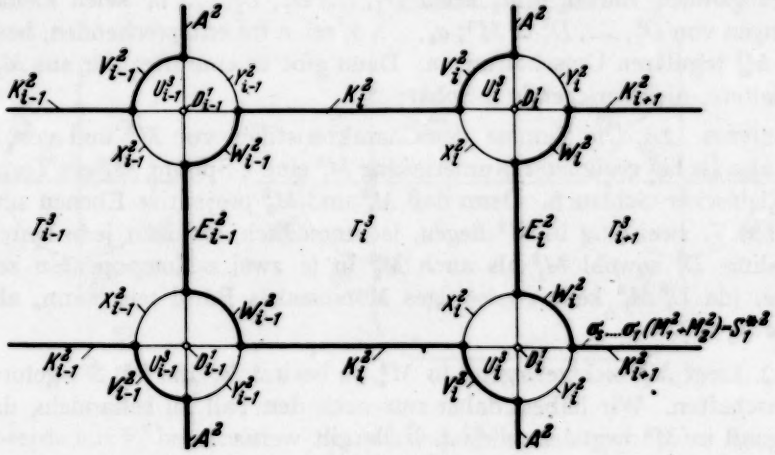

Fig. 10

$T^3_1+\dots+T^3_s+U^3_1+\dots+U^3_s$ ist eine 3-Mannigfaltigkeit M'^3. Dabei ist $M^3-M'^3$ zusammenhängend (da $\dot{M}'^3$ ein Torus bzw. Kleinscher Schlauch, also zusammenhängend, ist) und wird durch A^2 in zwei z-Komponenten Z^3_1 und Z^3_2 zerlegt. Denn wäre $M^3-(M'^3+A^2)$ zusammenhängend, so gäbe es in $M^3-(M'^3+A^2)$ einen Weg, der eine Seite von A^2 mit der anderen verbindet; dieser Weg verbände also auch in M^3 eine Seite von M^2_1 mit der anderen, ohne M^2_1 zu treffen; also wäre $M^3-M^2_1$ zusammenhängend und M^2_1 nichtzerlegend in M^3. — Bei geeigneter Numerierung von Z^3_1 und Z^3_2 liegen $K^2_1, K^2_3, \dots, K^2_{s-1}$ in $\dot{Z}^3_1$ und $K^2_2, K^2_4, \dots, K^2_s$ in $\dot{Z}^3_2$. Daraus folgt, daß $M^3-\sigma_s \dots \sigma_1(M^2_1+M^2_2)$ aus den beiden z-Komponenten $Z^3_1+\sum_{i \text{ gerade}}(\mathring{U}^3_i+X^2_i+Y^2_i+T^3_i)$ und $Z^3_2+\sum_{i \text{ ungerade}}(\mathring{U}^3_i+X^2_i+Y^3_i+T^3_i)$ besteht. Dieses steht im Widerspruch dazu, daß $\sigma_s \dots \sigma_1(M^2_1+M^2_2)$ (nach [5], Hauptsatz 4) eine zu S^2 ähnliche 2-Sphäre ist, also M^3 nicht zerlegt. Also ist die Annahme falsch, w.z.b.w.

2.5. Die 2-Sphäre S'^2 ist eine Sz-Komponente von $\sigma_{j-1}\bar{\sigma}_j(M^2_1+M^2_2)$, also irregulär aus $M^2_1+M^2_2$ abgeleitet, und besitzt damit die geforderten Eigenschaften. Damit ist der Beweis von Nr. 2 vollendet.

3. Nun läßt sich ebenso weiter schließen, wie in den Nrn. 3 und 4 des Beweises von Satz 5: Entweder ist die F-Zahl von S'^2 kleiner als f, oder sie ist gleich f und S'^2 enthält eine Falte. Daraus folgt (nach [5], Hauptsatz 1 und nach Satz 2 c), daß es eine Normalfläche S^{*2} gibt, die eine nichtzerlegende 2-Sphäre ist und eine F-Zahl kleiner als f hat. Liegt S^{*2} in keiner der Fundamentalklassen, so läßt sich wieder auf eine nichtzerlegende 2-Sphäre S^{**2} mit kleinerer F-Zahl schließen, usw., bis nach weniger als f Schritten eine nichtzerlegende 2-Sphäre erhalten wird, die in einer der Fundamentalklassen liegt. Damit ist Satz 6 bewiesen.

SATZ 7. *Der projektive Raum ist irreduzibel* [12]).

BEWEIS. Ein projektiver Raum P^3 läßt sich zerlegen in eine projektive Ebene P^2 und ein offenes 3-Element E^3 mit Randsingularitäten. Ist S^2 eine beliebige 2-Sphäre in P^3, so läßt sich durch kleine isotope Deformationen erreichen, daß S^2 mit P^2 nur geschlossene Durchdringungslinien gemeinsam hat. Da jede dieser Durchdringungslinien ein in P^2 liegendes 2-Element berandet, läßt sich S^2 durch ϱ-Operationen in ein Sphärensystem überführen, das in E^3 liegt, also nach [1] nur kompressible 2-Sphären enthält. Daraus folgt nach Satz 2 d), daß auch S^2 kompressibel in P^3 liegt. Also ist P^3 irreduzibel, w.z.b.w.

BESTIMMUNGSVERFAHREN 1. *Ist M^3 eine beliebig gegebene (zusammenhängende, berandete oder unberandete) 3-Mannigfaltigkeit, so lassen sich (in einer beschränkten Anzahl von Schritten) die Kreuzhaubenzahl k und die Henkelzahl h von M^3 bestimmen und es läßt sich eine kreuzhauben-, henkel- und randsphärenfreie 3-Mannigfaltigkeit M'^3 konstruieren, die aus M^3 durch k Kreuzhaubenreduktionen, h Henkelreduktionen und Ausfüllen der Randsphären hervorgeht.*

Dieses kann folgendermaßen geschehen: Es wird eine beliebige randtreue Zerlegung Γ_1 von M^3 gewählt und eine Normalzerlegung Θ_1 bezüglich Γ_1; dann werden (nach [5], Kap. II) die Fundamentalklassen bezüglich Θ_1 durch Konstruktion je eines Repräsentanten bestimmt. Ist keiner dieser Repräsentanten eine einseitig in M^3 liegende projektive Ebene, so ist M^3 nach Satz 5 kreuzhaubenfrei (d.h. es ist $k=0$). Ist dagegen (mindestens) einer der Repräsentanten — er werde mit P_1^2 bezeichnet — eine einseitige projektive Ebene, so wird eine kleine Umgebung K_1^3 von P_1^2 gewählt und eine 3-Mannigfaltigkeit $M_2^3 = M^3 - \mathring{K}_1^3$ gebildet. M_2^3 wird dann ebenso behandelt wie M^3: Es werden Zerlegung Γ_2, Normalzerlegung Θ_2 bezüglich Γ_2 und die Fundamentalklassen bezüglich Θ_2 konstruiert; befindet sich unter den Repräsentanten dieser Fundamentalklassen eine einseitige projektive Ebene P_2^2, so werden eine kleine Umgebung K_2^3 von P_2^2 und die 3-Mannigfaltigkeit $M_3^3 = M_2^3 - \mathring{K}_2^3$ hergestellt. So fortfahrend, erhalten wir nach einer auf Grund des Endlichkeitssatzes 1 beschränkten Anzahl von Schritten eine kreuzhaubenfreie 3-Mannigfaltigkeit $M_{k+1}^3 = M_k^3 - \mathring{K}_k^3$.

[12]) Die Irreduzibilität aller Linsenräume wird in einer späteren Arbeit bewiesen werden.

Gibt es nun unter den Repräsentanten der Fundamentalklassen bezüglich Θ_{k+1} keine 2-Sphäre, die nichtzerlegend in M^3_{k+1} liegt, so ist M^3_{k+1} nach Satz 6 henkelfrei. Ist dagegen (mindestens) einer der Repräsentanten — er werde mit S^2_1 bezeichnet — eine nichtzerlegende 2-Sphäre in M^3_{k+1}, so wird eine kleine Umgebung H^3_1 von S^2_1 in M^3_{k+1} gewählt und die 3-Mannigfaltigkeit $M^3_{k+2} = M^3_{k+1} - \mathring{H}^3_1$ gebildet. Mit M^3_{k+2} verfahren wir ebenso weiter wie mit M^3_{k+1}, usw., und erhalten nach einer beschränkten Anzahl von Schritten eine kreuzhauben- und henkelfreie 3-Mannigfaltigkeit $M^3_{k+h+1} = M^3_{k+h} - \mathring{H}^3_h$. Schließlich füllen wir die Randsphären von M^3_{k+h+1} durch 3-Elemente aus und erhalten dadurch eine 3-Mannigfaltigkeit M'^3, die aus M^3 durch k Kreuzhaubenreduktionen, h Henkelreduktionen und Ausfüllen der Randsphären hervorgeht; sie ist kreuzhauben-, henkel- und randsphärenfrei.

k und h sind die Kreuzhauben- bzw. Henkelzahl von M^3. Denn in M^3_{k+h+1} existiert nach der Folgerung aus Endlichkeitssatz 1 ein vollständiges Sphärensystem $T^2_1, \ldots, T^2_m$; es zerlegt M^3_{k+h+1} in $m+1$ z-Komponenten, wobei sich keine von diesen durch 3-Elemente zu einem projektiven Raum ergänzen läßt. $T^2_1, \ldots, T^2_m, \dot{K}^3_1, \ldots, \dot{K}^3_k, \dot{H}^3_1, \ldots, \dot{H}^3_h$ bilden ein vollständiges Sphärensystem in M^3 (da sich die H^3_i durch 3-Elemente zu 3-Sphären ergänzen lassen und die K^3_j zu projektiven Räumen, die nach Satz 7 irreduzibel sind); also besitzt die Äquivalenzklasse $\mathfrak{M}^3$ von M^3 nach Eindeutigkeitssatz 1 h Henkel und k mal den projektiven Raum als irreduzible Erzeugende.

Damit sind die Zahlen k und h und die 3-Mannigfaltigkeit M'^3 erhalten.

FOLGERUNG. *Das in* [5], *Kap.* V.4 *angegebene Bestimmungsverfahren für Bänder minimaler Charakteristik mit vorgegebener Randlinie und das Kreislinienkriterium* ([5], *Kap.* V.5) *lassen sich für beliebige 3-Mannigfaltigkeiten verallgemeinern:*

Ist M^3 eine beliebige (zusammenhängende, berandete oder unberandete) 3-Mannigfaltigkeit, in deren Rand eine 1-Sphäre Y^1 vorgegeben ist, so können wir M^3 nach dem eben dargestellten Bestimmungsverfahren durch Kreuzhauben- und Henkelreduktionen in eine kreuzhauben- und henkelfreie 3-Mannigfaltigkeit M'^3 überführen. Dann läßt sich in M'^3 nach dem Bestimmungsverfahren [5], Kap. V.4 ein in Y^1 eingespanntes Band minimaler Charakteristik konstruieren (sofern es überhaupt in Y^1 eingespannte Bänder in M'^3 gibt); es läßt sich in M'^3 so isotop deformieren, daß es die Ränder der reduzierten Henkel und Kreuzhauben nicht trifft, also auch in M^3 liegt. Dieses Band ist auch von minimaler Charakteristik in M^3. Denn jedes in M^3 liegende (in Y^1 eingespannte) Band läßt sich durch ϱ-Operationen so abändern, daß ein Band von nicht größerer Charakteristik entsteht, das die Ränder der Kreuzhauben und Henkel nicht trifft, also auch in M'^3 liegt.

Ist in M^3 eine Knotenlinie (1-Sphäre) K^1 vorgegeben, so kann man ebenso vorgehen, wie beim Kreislinienkriterium [5], Kap. V.5 beschrieben, und hat dabei lediglich statt des in [5], Kap. V.4 dargestellten Bestimmungsverfahrens das oben angegebene verallgemeinerte Bestimmungsverfahren anzuwenden.

§ 5. Bestimmung eines vollständigen Sphärensystems

Sind an einer gegebenen 3-Mannigfaltigkeit M^3 alle Kreuzhauben und Henkel reduziert worden, so kommt es nun darauf an, ein vollständiges System von 2-Sphären zu bestimmen, um M^3 in irreduzible Erzeugende aufzuspalten. Da wir kein Äquivalenzkriterium für die 3-Sphäre besitzen, können wir diese Aufspaltung, im Gegensatz zur Reduktion der Kreuzhauben und Henkel, nicht successive in endlich vielen Schritten durchführen, sondern müssen uns mit *einem* Schritt ein vollständiges Sphärensystem verschaffen. Hierdurch werden die Untersuchungen komplizierter als in §4, weil Normalflächen behandelt werden müssen, die nicht zusammenhängend sind.

HILFSSATZ 2. *M^3 sei eine kreuzhauben-, henkel- und randsphärenfreie, sonst beliebige (zusammenhängende, berandete oder unberandete) 3-Mannigfaltigkeit. Θ sei eine Normalzerlegung bezüglich einer beliebigen randtreuen Zerlegung Γ von M^3 in offene Raumelemente. Weiter sei in M^3 ein vollständiges Sphärensystem $S_1^2, \ldots, S_m^2$ gegeben. Dabei seien $S_1^2, \ldots, S_m^2$ Normalflächen bezüglich Θ mit den P-Zahlenvektoren $\mathfrak{x}_1, \ldots, \mathfrak{x}_m$. Dann folgt:*

Sind nicht alle Vektoren $\mathfrak{x}_1, \ldots, \mathfrak{x}_m$ Fundamentallösungen der P-Gleichungen so existiert in M^3 ein vollständiges Sphärensystem $S_1'^2, \ldots, S_{m'}'^2$, so daß $S_1'^2, \ldots, S_{m'}'^2$ ebenfalls Normalflächen sind und die F-Zahl f' von $S_1'^2 + \cdots + S_{m'}'^2$ kleiner ist als die F-Zahl f von $S_1^2 + \cdots + S_m^2$.

BEWEIS. Wir können voraussetzen, daß $\mathfrak{x}_1$ keine Fundamentallösung sei.

1. Zunächst ergibt sich ebenso, wie beim Beweis von Satz 6, Nr. 2, 2.1 und 2.3 (s. Fig. 10): Es gibt ein reduziertes, S_1^2 erzeugendes Normalflächenpaar (M_1^2, M_2^2). Die Durchdringungslinien von $M_1^2 + M_2^2$ seien $D_1^1, \ldots, D_s^1$; $U_1^3, \ldots, U_s^3$ seien (im Verhältnis zu $M_1^2 + M_2^2 | \Theta$ kleine) Umgebungen von $D_1^1, \ldots, D_s^1$ in M^3; $\sigma_1, \ldots, \sigma_s$ seien die entsprechenden regulären Umschaltungen. Bei geeigneter Numerierung ist M_1^2 eine 2-Sphäre und M_2^2 ein Torus oder ein Kleinscher Schlauch; M_1^2 liegt n.V. zerlegend in M^3; s ist daher eine gerade Zahl. Für die z-Komponenten von $M_1^2 - (U_1^3 + \cdots + U_s^3)$, $M_2^2 - (U_1^3 + \cdots + U_s^3)$ und $\dot{U}_i^3 - (M_1^2 + M_2^2)$ führen wir dieselben Bezeichnungen ein, wie beim Beweis von Satz 6 (s. Fig. 10). Die 2-Sphäre $E_{i-1}^2 + \overline{W}_{i-1}^2 + K_i^2 + \overline{X}_i^2 + E_i^2$ zerlegt M^3 n.V. in zwei z-Komponenten, deren eine — sie werde mit T_i^3 bezeichnet — mit A^2 punktfremd ist. Die zu S_1^2 ähnliche 2-Sphäre $\sigma_s \ldots \sigma_1 (M_1^2 + M_2^2)$ bezeichnen wir zur Abkürzung mit S_1^{*2}.

2. Nun existieren in M^3 mit S_1^{*2} punktfremde 2-Sphären $S_2^{*2}, \ldots, S_{m^*}^{*2}$, so daß folgendes gilt:

a) $S_1^{*2} + \cdots + S_{m^*}^{*2}$ ist ein vollständiges Sphärensystem in M^3.

b) $S_2^{*2} + \cdots + S_{m^*}^{*2}$ ist eine Halbnormalfläche bezüglich Θ, deren F-Zahl nicht größer ist als die von $S_2^2 + \cdots + S_m^2$.

c) Der Durchschnitt $(S_2^{*2} + \cdots + S_{m^*}^{*2})\, U_i^3$ ist entweder leer, oder besteht aus Kreisringen $R_{i1}^2, \ldots, R_{ii_i}^2$, deren Inneres in $\mathring{U}_i^3$ liegt, und deren jeder eine

in X_i^2 und eine in Y_i^2 liegende Randlinie hat, so daß diese Randlinien X_i^2 und Y_i^2 in (offene) Kreisringe zerlegen (für alle $i=1, \ldots, s$).

BEWEIS. Zunächst gibt es ein zu $S_2^2+\cdots+S_m^2$ bezüglich Θ ähnliches System von 2-Sphären $S_{(1)2}^2, \ldots, S_{(1)m}^2$, das mit S_1^{*2} punktfremd ist. (Ein solches Sphärensystem entsteht z. B. als Bild von $S_2^2+\cdots+S_m^2$ bei einer semilinearen Abbildung von M^3 auf sich selbst, die S_1^2 in S_1^{*2} überführt und die Elemente von Θ jeweils in sich selbst abbildet.) $S_{(1)2}^2, \ldots, S_{(1)m}^2$ lassen sich durch kleine isotope Deformationen in bezüglich Θ ähnliche 2-Sphären $S_{(2)2}^2, \ldots, S_{(2)m}^2$ überführen, die mit den Kreisringen X_i^2 höchstens Durchdringungslinien gemeinsam haben. Befinden sich unter diesen Durchdringungslinien solche, die in den X_i^2 liegende 2-Elemente beranden, so lassen sich die Durchdringungslinien dieser Art successive durch ϱ-Operationen beseitigen, deren Kerne im Verhältnis zu Θ kleine Umgebungen von 2-Elementen aus den X_i^2 sind, so daß die F-Zahl des Sphärensystems nicht verändert wird und die Halbnormalflächeneigenschaft erhalten bleibt. Die so erhaltenen 2-Sphären $S_{(3)2}^2, \ldots, S_{(3)m}^2$ haben mit den X_i^2 höchstens solche Durchdringungslinien gemeinsam, die die X_i^2 in Kreisringe zerlegen, und bilden nach Satz 3 zusammen mit S_1^{*2} ein vollständiges Sphärensystem in M^3. Nun sei in jeder der Umgebungen U_i^2 ein (im Verhältnis zu $S_{(3)2}^2+\cdots+S_{(3)m}^2|\Theta$) nahe zu $V_i^2+\bar{X}_i^2+W_i^2$ benachbarter Kreisring Y_i^{*2} ausgezeichnet (s. Fig. 11 und Fig. 10); der von $Y_i^{*2}+V_i^2+\bar{X}_i^2+W_i^2$ berandete, in U_i^3 liegende Vollring werde mit U_i^{*3} bezeichnet. Die 2-Sphären $S_{(3)2}^2, \ldots, S_{(3)m}^2$ schneiden U_i^{*3} in Kreisringen $R_{i1}^{*2}, \ldots, R_{it_i}^{*2}$ ($t_i=0, 1, 2, \ldots$), deren Inneres in $\mathring{U}_i^{*3}$ liegt, und deren jeder eine in X_i^2 und eine in Y_i^{*2} liegende Randlinie besitzt. Nun gibt es eine semilineare Abbildung ζ von M^3 auf sich selbst, durch die U_i^{*3} in U_i^3 (für alle $i=1, \ldots, s$) übergeht, so daß V_i^2, W_i^2 und X_i^2 in sich selbst abgebildet werden und Y_i^{*2} in Y_i^2 übergeht. Die Abbildung ζ kann ferner so gewählt werden, daß S_1^{*2} fest bleibt und die einzelnen Elemente von Θ jeweils in sich selbst abgebildet werden. Durch ζ geht $S_{(3)2}^2+\cdots+S_{(3)m}^2$ in ein Sphärensystem $S_2^{*2}+\cdots+S_m^{*2}$ über, das die geforderten Eigenschaften hat. Damit ist Nr. 2 bewiesen.

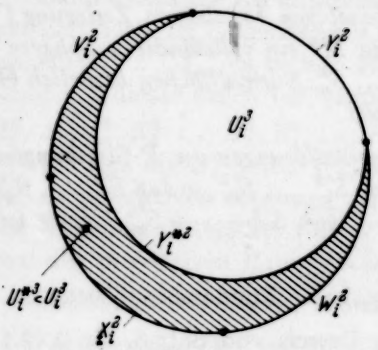

Fig. 11

3. Die Anzahl der Kreisringe R_{ij}^2 bezeichnen wir mit $r\left(=\sum_{i=1}^{s} t_i\right)$ und nennen sie kurz die *R-Zahl* des Sphärensystems $S_1^{*2}+\cdots+S_m^{*2}$. Wir wollen nun beweisen, daß sich aus $S_1^{*2}+\cdots+S_m^{*2}$ ein neues, ebenfalls vollständiges Sphärensystem in M^3 ableiten läßt, das ebenfalls eine Halbnormalfläche ist und entweder eine F-Zahl kleiner als f hat, oder die F-Zahl f besitzt und eine Falte enthält. Daraus folgt dann Hilfssatz 2 nach [5], Hauptsatz 1 und nach Satz 3.

Den Beweis führen wir so, daß wir aus $S_1^{*2}+\cdots+S_{m^*}^{*2}$ zunächst ein vollständiges Sphärensystem $S_1^{(2)2}+\cdots+S_{m^*}^{(2)2}$ ableiten, das eine Halbnormalfläche ist, so daß einer der beiden folgenden Fälle vorliegt:

I. Die F-Zahl $f^{(2)}$ von $S_1^{(2)2}+\cdots+S_{m^*}^{(2)2}$ ist entweder kleiner als f, oder es ist $f^{(2)}=f$ und $S_1^{(2)2}+\cdots+S_{m^*}^{(2)2}$ enthält eine Falte; d.h. das Sphärensystem hat bereits die geforderten Eigenschaften.

II. Es ist $f^{(2)}=f$ und es gilt:

II.1. $S_1^{(2)2}$ ist gleich S_1^{*2},

II.2. der Durchschnitt $(S_2^{(2)2}+\cdots+S_{m^*}^{(2)2})\,U_i^3$ ist entweder leer oder besteht aus Kreisringen $R_{i1}^{(2)2},\ldots,R_{it_i^{(2)}}^{(2)2}$ mit den entsprechenden Eigenschaften, wie sie in Nr. 2 Punkt c) aufgeführt wurden,

II.3. die R-Zahl $r^{(2)}=\sum_{i=1}^{s} t_i^{(2)}$ ist kleiner als r.

Wenn der Fall II vorliegt, leiten wir ein weiteres Sphärensystem $S_1^{(3)2},\ldots,S_{m^*}^{(3)2}$ mit analogen Eigenschaften ab und fahren so fort, bis wir nach höchstens $r+1$ Schritten zu einem Sphärensystem kommen, für das der angestrebte Fall I vorliegt.

Wir haben also noch zu beweisen: Ist das Sphärensystem $S_2^{(x)2}+\cdots+S_{m^*}^{(x)2}$ gegeben, und liegt der Fall II vor, so existiert ein Sphärensystem $S_1^{(x+1)2},\ldots,$

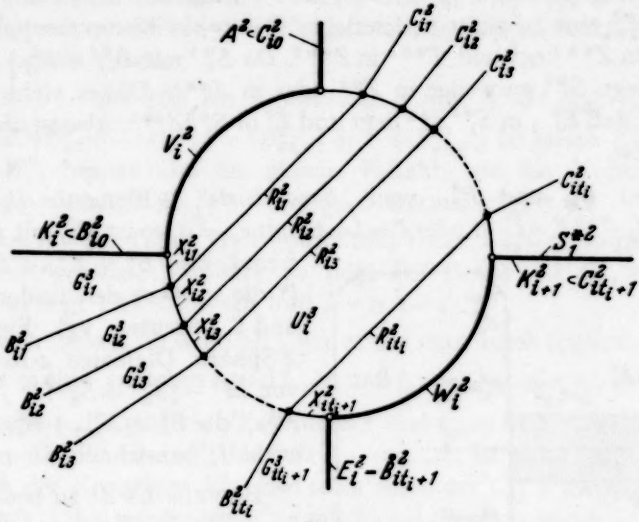

Fig. 12

$S_{m^*}^{(x+1)2}$ für das I oder II (mit $r^{(x+1)}<r^{(x)}$) gilt. — Die Kreisringe $R_{i1}^{(x)2},\ldots,R_{it_i^{(x)}}^{(x)2}$ bezeichnen wir der Einfachheit halber, und da Verwechslungen ausgeschlossen sind, wieder mit $R_{i1}^2,\ldots,R_{it_i}^2$. Die Kreisringe, in die X_i^2 durch die $\dot{R}_{ij}^2$ zerlegt wird, bezeichnen wir mit $X_{i1}^2,\ldots,X_{it_i+1}^2$. Dann gilt weiter:

3.1. Bei geeigneter Numerierung liegen die Kreisringe X_{ik}^2 und R_{ij}^2 in der Reihenfolge ihrer Numerierung in X_i^2 bzw. U_i^3 (s. Fig. 12), d.h. $\dot{X}_{i1}^2$ enthält

eine 1-Sphäre aus $\dot{K}_i^2$, die 1-Sphäre $\dot{R}_{ij}^2 X_i^2$ liegt in $\dot{X}_{ij}^2$ und in $\dot{X}_{ij+1}^2$ (für alle $j=1,\ldots,t_i$). — Der Kreisring R_{ij}^2 liegt in einer 2-Sphäre $S_{a(i,j)}^{(s)2}$ mit $2 \leqq a(i,j) \leqq m^*$ und $S_{a(i,j)}^{(s)2} - \mathring{R}_{ij}^2$ besteht aus zwei 2-Elementen, die wir mit B_{ij}^2 und C_{ij}^2 bezeichnen, so daß $\dot{B}_{ij}^2 = \dot{R}_{ij}^2 X_i^2$ und $\dot{C}_{ij}^2 = \dot{R}_{ij}^2 Y_i^2$ ist. Der Einheitlichkeit halber bezeichnen wir das 2-Element $\bar{K}_i^2 + W_{i-1}^2 + \bar{E}_{i-1}^2$ mit B_{i0}^2 und $\bar{E}_i^2$ mit $B_{it_i+1}^2$, entsprechend $S_1^{*2} - (B_{i0}^2 + V_i^2)$ mit C_{i0}^2 und $S_1^{*2} - (B_{it_i+1}^2 + W_i^2)$ mit $C_{it_i+1}^2$ (vgl. Fig. 12). Die F-Zahl von B_{ij}^2 nennen wir f_{ij} und die R-Zahl r_{ij}. (Die einzelnen B_{ij}^2 brauchen keineswegs ganz in $\bar{T}_i^3$ zu liegen, insbesondere kann $B_{ij_1}^2$ in $B_{ij_2}^2$ liegen.) Entsprechend seien f_{i0} bzw. f_{it_i+1} und r_{i0} bzw. r_{it_i+1} die F- und R-Zahlen von B_{i0}^2 bzw. $B_{it_i+1}^2$ ($r_{i0} = r_{it_i+1} = 0$). Die kleinste der Zahlen $f_{i1}, \ldots, f_{it_i}$ nennen wir g_i.

3.2. Wir beweisen zunächst zwei Eigenschaften der 2-Elemente B_{ij}^2 und C_{ij}^2:

3.2.1. Jedes 2-Element C_{ij}^2 ($i=1,\ldots,s$; $j=1,\ldots,t_i$) enthält mindestens eines der $B_{i1}^2, \ldots, B_{it_i}^2$, besitzt also eine F-Zahl, die nicht kleiner als g_i ist.

BEWEIS. Aus der Annahme, C_{ij}^2 sei mit T_i^3 punktfremd, folgt: Wir können in $\bar{T}_i^3$ ein beliebiges 2-Element B_{ij}^{*2} mit $\dot{B}_{ij}^{*2} = \dot{B}_{ij}^2$ wählen, dessen Inneres in $\mathring{T}_i^3$ liegt. Dann besteht $\bar{T}_i^3 - B_{ij}^{*2}$ aus zwei z-Komponenten T^{*3} und T^{**3}, so daß (vgl. Fig. 10 und Fig. 12) E_{i-1}^2 in $\dot{T}^{*3}$ liegt und E_i^2 in $\dot{T}^{**3}$. Nun ist $B_{ij}^{*2} + R_{ij}^2 + C_{ij}^2$ eine 2-Sphäre und zerlegt M^3 in zwei z-Komponenten Z^{*3}, Z^{**3}, so daß T^{*3} in Z^{*3} liegt und T^{**3} in Z^{**3}. Da S_1^{*2} mit $B_{ij}^{*2} + R_{ij}^2 + C_{ij}^2$ punktfremd ist, liegt S_1^{*2} entweder in Z^{*3} oder in Z^{**3}. Dieses steht im Widerspruch dazu, daß E_{i-1}^2 in $S_1^{*2} Z^{*3}$ liegt und E_i^2 in $S_1^{*2} Z^{**3}$. Also ist die Annahme falsch, w.z.b.w.

3.2.2. Sind B_{kl}^2 und $B_{kl'}^2$ zwei *„benachbarte"* 2-Elemente ($k=1,\ldots,s$; $l=0,1,\ldots,t_k+1$; $l'=l-1$ oder $l'=l+1$) mit $r_{kl}=0$, so ist B_{kl}^2 mit $B_{kl'}^2$ punktfremd, also $B_{kl}^2 + X_{kl^*}^2 + B_{kl'}^2$ (wobei l^* die größere der beiden Zahlen l und l' bedeutet, vgl. Fig. 12) eine 2-Sphäre. Diejenige z-Komponente von $M^3 - (S_1^{*2} + S_2^{(s)2} + \cdots + S_{m^*}^{(s)2} + X_{kl^*}^2)$, die $B_{kl}^2 + X_{kl^*}^2 + B_{kl'}^2$ im Rande enthält, bezeichnen wir mit $G_{kl^*}^3$.

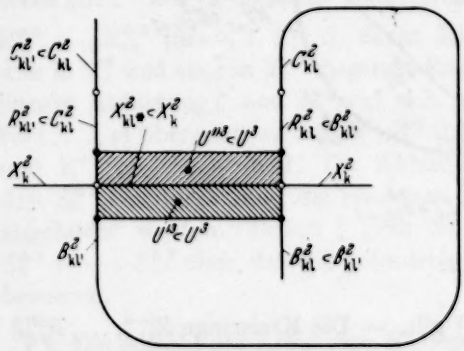

Fig. 13

BEWEIS. Es ist zu beweisen, daß B_{kl}^2 nicht in $B_{kl'}^2$ liegt. Dieses gilt offenbar, wenn $l^*=1$ oder $l^*=t_k+1$ ist. Wir können daher $l^* \neq 1$, t_k+1 voraussetzen. Aus der Annahme B_{kl}^2 läge in $B_{kl'}^2$, folgt: Die 2-Sphäre $S_{a(k,l)}^{(s)2}$ ist gleich $B_{kl'}^2 + R_{kl'}^2 + C_{kl'}^2$ und zerlegt M^3 in zwei z-Komponenten, deren eine — sie werde mit Z^3 bezeichnet — den Kreisring $X_{kl^*}^2$ enthält. Wir betrachten eine kleine Umgebung U^3 von $\bar{X}_{kl^*}^2$ in $\bar{Z}^3$ (s. Fig. 13). $U^3 - \bar{X}_{kl^*}^2$ besteht aus zwei z-Komponenten U'^3 und U''^3, so daß U'^3 in T_k^3 liegt und U''^3 in U_k^3.

Nun ist $B^2_{kl}+X^2_{kl\bullet}+R^2_{kl'}+C^2_{kl'}$ eine 2-Sphäre, die M^3 in zwei z-Komponenten Z'^3 und Z''^3 zerlegt, so daß $\overset{\circ}{U}{}'^3$ in Z'^3 liegt und $\overset{\circ}{U}{}''^3$ in Z''^3. Der offene Kreisring $(B^2_{kl'}-\overset{\circ}{B}{}^2_{kl})$ ist mit $B^2_{kl}+X^2_{kl\bullet}+R^2_{kl'}+C^2_{kl'}$ punktfremd, liegt also entweder in Z'^3 oder in Z''^3. Dieses steht im Widerspruch dazu, daß $\dot{U}'^3\overset{\circ}{B}{}^2_{kl'}$ in $Z'^3(B^2_{kl'}-\overset{\circ}{B}{}^2_{kl})$ liegt und $\dot{U}''^3\overset{\circ}{R}{}^2_{kl}$ in $Z''^3(B^2_{kl'}-\overset{\circ}{B}{}^2_{kl})$ (vgl. Fig. 13). Also ist die Annahme falsch, w.z.b.w.

3.3. Es liegt stets einer der folgenden Fälle vor:

A. Es gibt zwei benachbarte 2-Elemente B^2_{kl} und $B^2_{kl'}$ mit $r_{kl}=0$, wie unter 3.2.2 betrachtet, so daß $f_{kl}\leqq g_k$ oder $t_k=0$ ist ($f_{kl}<g_k$ ist nur dann möglich, wenn $l=0$ oder $l=t_k+1$ ist) und so daß einer der beiden Fälle vorliegt:

A.1. $G^3_{kl\bullet}$ ist keine gelochte 3-Sphäre, und es ist

α) $l'\neq 0$ und $l'\neq t_k+1$ *oder*

β) die F-Zahl von $C^2_{kl'}$ ist größer als f_{kl}.

A.2. $G^3_{kl\bullet}$ ist eine gelochte 3-Sphäre, und es ist entweder

α) $f_{kl'}>f_{kl}$ *oder*

β) $f_{kl'}=f_{kl}$ und $r_{kl'}\neq 0$.

B. Es gibt ein k (zwischen 1 und s) mit $f_{k0}>f_{k1}=f_{k2}=\cdots=f_{kt_k+1}$ und $r_{k1}=\cdots=r_{kt_k}=0$, so daß G^3_{k1} keine gelochte 3-Sphäre ist und $G^3_{k2},\ldots,G^3_{kt_k+1}$ gelochte 3-Sphären sind.

Beweis. 3.3.1. Es gibt ein k (zwischen 1 und s) mit $f_{k0}>f_{kt_k+1}$.

Beweis. $(B^2_{10}+\cdots+B^2_{s0})-(B^2_{1t_1+1}+\cdots+B^2_{st_s+1})$ ist gleich $\bar{K}^2_1+\cdots+\bar{K}^2_s+W^2_1+\cdots+W^2_s$, besitzt also die gleiche F-Zahl, wie die Normalfläche M^2_2. Also ist $f_{10}+\cdots+f_{s0}>f_{1t_1+1}+\cdots+f_{st_s+1}$. Daraus folgt 3.3.1.

3.3.2. Ist für den (nach 3.3.1 gewählten) Index k $t_k=0$, so gilt 3.3. Denn ist G^3_{k1} keine gelochte 3-Sphäre, so liegt Fall B vor; ist G^3_{k1} eine gelochte 3-Sphäre, so liegt Fall A.2.α vor (mit $l'=0$, $l=1$).

3.3.3. Den Beweis für den Fall $t_k\neq 0$ erhält man durch folgende Gliederung:

a) Gibt es ein j' zwischen 1 und t_k, so daß $f_{kj'}>g_k$ ist *oder* $r_{kj'}\neq 0$, so gilt 3.3.

Beweis. a 1. Es gibt ein j'' zwischen 1 und t_k, so daß $f_{kj''}=g_k$ und $r_{kj''}=0$ ist. Denn n. V. gibt es ein j''', so daß $f_{kj'''}=g_k$ ist; ist nun $r_{kj'''}\neq 0$, so enthält $B^2_{kj'''}$ einen der Kreisringe R^2_{kj}, also auch eines der C^2_{kj} (j zwischen 1 und t_k), und damit (s. 3.2.1) wieder ein $B^2_{kj^{IV}}$ mit $f_{kj^{IV}}=g_k$ und $r_{kj^{IV}}<r_{kj'''}$; ist $r_{kj^{IV}}\neq 0$, so kann man weiter auf ein j^{V} schließen, usw., bis man nach höchstens $r_{kj'''}$ Schritten zu einem $j^{(w)}=j''$ mit $f_{kj^{(w)}}=g_k$, $r_{kj^{(w)}}=0$ gelangt.

a 2. Es gibt nun zwei Indizes l und l', so daß entweder $j'\leqq l'=l-1<j''$ oder $j'\geqq l'=l+1>j''$ ist, so daß $f_{kl}=g_k$, $r_{kl}=0$ ist, und so daß $f_{kl'}>g_k$ oder $r_{kl'}\neq 0$ ist. Also liegt Fall A.1.α oder Fall A.2 (α oder β) vor. Damit ist a) bewiesen.

b) Ist $f_{k1}=\cdots=f_{kt_k}=g_k$ und $r_{k1}=\cdots=r_{kt_k}=0$, so gilt 3.3 ebenfalls.

BEWEIS. b 1. Gibt es ein l^* zwischen 2 und t_k, so daß $G^3_{kl^*}$ keine gelochte 3-Sphäre ist, so liegt Fall A.1.α vor.

b 2. Sind $G^3_{k2}, \ldots, G^3_{kt_k}$ gelochte 3-Sphären, so gilt 3.3.

BEWEIS. b 2.1. Ist $g_k > f_{kt_k+1}$, so liegt (mit $l=t_k+1$, $l'=t_k$) Fall A.1.α oder Fall A.2.α vor.

b 2.2. Ist $g_k = f_{kt_k+1}$, so gilt:

b 2.2.1. Ist $G^3_{kt_k+1}$ keine gelochte 3-Sphäre, so liegt (mit $l=t_k$, $l'=t_k+1$) Fall A.1.β vor (da $C^2_{kt_k+1}$ das 2-Element B^2_{k0} enthält, also eine F-Zahl größer als g_k hat).

b 2.2.2. Ist $G^3_{kt_k+1}$ eine gelochte 3-Sphäre, so liegt entweder Fall B vor, oder (mit $l=1$, $l'=0$) Fall A.2.α.

b 2.3. Ist $g_k < f_{kt_k+1}$, so liegt (mit $l=t_k, l'=t_k+1$) Fall A.1.β oder A.2.α vor.

Damit ist der Beweis für $t_k \neq 0$ vollendet und 3.3 vollständig bewiesen.

3.4. Zunächst machen wir eine Bemerkung über Abänderungen des Sphärensystems $S_1^{(z)2}+\cdots+S_{m^*}^{(z)2}$: E^2 sei ein 2-Element, dessen Rand in einer 2-Sphäre $S_v^{(z)2}$ (v zwischen 1 und m^*) liegt, und das im übrigen mit dem Sphärensystem punktfremd ist. Dann liegt $\mathring{E}^2$ in einer z-Komponente Z^3 von $M^3-(S_1^{(z)2}+\cdots+S_{m^*}^{(z)2})$ und zerlegt diese in zwei z-Komponenten Z'^3, Z''^3. Durch $\dot{E}^2$ wird $S_v^{(z)2}$ in zwei 2-Elemente E'^2 und E''^2 zerlegt, so daß E'^2 in $\dot{Z}'^3$ liegt und E''^2 in $\dot{Z}''^3$. Nun ist (mindestens) eine der Komponenten Z'^3, Z''^3 eine gelochte 3-Sphäre. (Denn anderenfalls ließe sich Z^3 durch eine nahe zu $E^2+E'^2$ benachbarte 2-Sphäre in zwei z-Komponenten zerlegen, deren keine eine gelochte 3-Sphäre ist; also wäre das Sphärensystem nicht vollständig.) Ist also Z'^3 eine gelochte 3-Sphäre, so kann man das 2-Element E'^2 durch E^2 ersetzen, und das Sphärensystem $S_1^{(z)2}+\cdots+S_{m^*}^{(z)2}$ geht dadurch in ein Sphärensystem über, das ebenfalls vollständig ist.

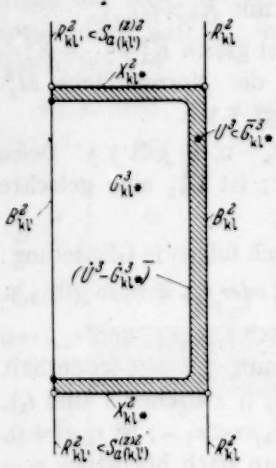
Fig. 14

3.5. Wir erhalten nun das Sphärensystem $S_1^{(z+1)2}+\cdots+S_{m^*}^{(z+1)2}$ durch folgende Abänderungen von $S_1^{(z)2}+\cdots+S_{m^*}^{(z)2}$:

3.5.1. Liegt der Fall A.1 (s. 3.3) vor, so wählen wir in $\overline{G}^3_{kl^*}$ eine (im Verhältnis zu $\Theta|U_k^3$ kleine) Umgebung U^3 von $B^2_{kl}+\overline{X}^2_{kl^*}$, s. Fig. 14. Dann ersetzen wir das 2-Element $S^{(z)2}_{a(k,l')}-(B^2_{kl'}-U^3)$ durch das zu $B^2_{kl}+\overline{X}^2_{kl^*}$ benachbarte 2-Element $(\dot{U}^3-\dot{G}^3_{kl^*})^-$ und erhalten so das gesuchte Sphärensystem $S_1^{(z+1)2}+\cdots+S_{m^*}^{(z+1)2}$. Die Numerierung der erhaltenen 2-Sphären wählen wir so, daß die nicht veränderten 2-Sphären jeweils ihren Index behalten. Wir überzeugen uns davon, daß dieses Sphärensystem die geforderten Eigenschaften besitzt:

a) Da die durchgeführte Abänderung von der in 3.4 beschriebenen Form ist, ist das Sphärensystem vollständig; ferner ist es eine Halbnormalfläche (da U^3 klein im Verhältnis zu Θ ist).

b) Im Falle A.1.α ist $S_1^{(z+1)2} = S_1^{(z)2} = S_1^{*2}$, und $S_2^{(z+1)2} + \cdots + S_{m^*}^{(z+1)2}$ schneidet, ebenso wie $S_1^{(z)2} + \cdots + S_{m^*}^{(z)2}$, die Umgebungen U_i^3 in Kreisringen. Dabei ist die R-Zahl von $S_{a(k,l')}^{(z+1)2}$ gleich der von $B_{kl'}^2$, also kleiner als die von $S_{a(k,l')}^{(z)2}$. Also ist $r^{(z+1)} < r^{(z)}$. Die F-Zahl von $S_{a(k,l')}^{(z+1)2}$ ist gleich $f_{kl} + f_{kl'}$, also nach 3.2.1 nicht größer als die von $S_{a(k,l')}^{(z)2}$. Also ist $f^{(z+1)} \leqq f^{(z)}$. Damit sind alle Bedingungen für das Vorliegen des Falles I oder II erfüllt.

c) Im Falle A.1.β ist die F-Zahl $f_{kl} + f_{kl'}$ von $S_{a(k,l')}^{(z+1)2}$ kleiner als die F-Zahl von $S_{a(k,l')}^{(z)2}$. Also ist $f^{(z+1)} < f^{(z)}$, und es liegt der Fall I vor.

3.5.2. Liegt der Fall A.2 vor, so wählen wir (wie in 3.5.1) eine kleine Umgebung U^3 von $B_{kl}^2 + \overline{X}_{kl^*}^2$ in $\overline{G}_{kl^*}^3$ (vgl. Fig. 14). Dann ersetzen wir $B_{kl'}^2 - U^3$ durch $\dot{U}^3 - G_{kl^*}^3$. Das so entstehende Sphärensystem $S_1^{(z+1)2} + \cdots + S_{m^*}^{(z+1)2}$ hat (bei entsprechender Numerierung) die geforderten Eigenschaften, denn im Falle A.2.α ist $f^{(z+1)} < f^{(z)}$, und im Falle A.2.β ist $f^{(z+1)} = f^{(z)}$, $r^{(z+1)} < r^{(z)}$, sowie $S_1^{(z+1)2} = S_1^{(z)2} = S_1^{*2}$ [da bei $r_{kl'} \neq 0$ stets $a(k,l') \neq 1$ ist].

3.5.3. Liegt der Fall B vor, so wählen wir in $\overline{T}_k^3$ ein (im Verhältnis zu $S_1^{(z)2} + \cdots + S_{m^*}^{(z)2} | \Theta$) nahe zu $B_{kt_k+1}^2 + X_{kt_k+1}^2$ benachbartes 2-Element, das wir

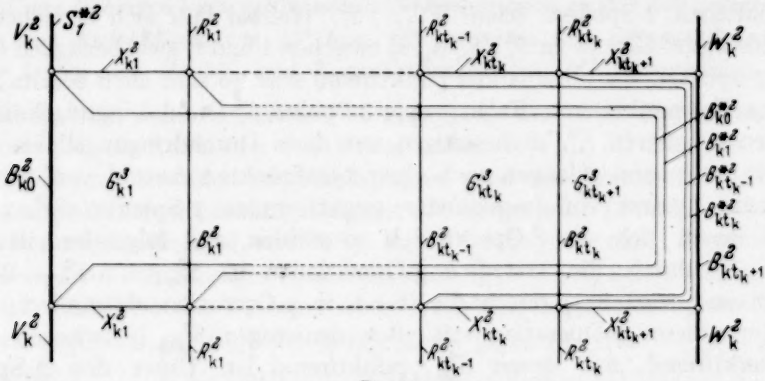

Fig. 15

$B_{kt_k}^{*2}$ nennen, und ersetzen $B_{kt_k}^2$ durch $B_{kt_k}^{*2}$. Dann wählen wir ein zu $B_{kt_k}^{*2} + X_{kt_k}^2$ nahe benachbartes 2-Element $B_{kt_k-1}^{*2}$ und ersetzen $B_{kt_k-1}^2$ durch $B_{kt_k-1}^{*2}$, und so fort, bis wir B_{k1}^2 durch B_{k1}^{*2} ersetzt haben, s. Fig. 15. Schließlich wählen wir ein zu $B_{k1}^{*2} + X_{k1}^2$ benachbartes 2-Element B_{k0}^{*2} und ersetzen $S_1^{*2} - B_{k0}^2$ durch B_{k0}^{*2}. (Dieses ist eine Folge von Abänderungen der in 3.4 beschriebenen Art.) Die entstehenden 2-Sphären numerieren wir so, daß $B_{k0}^2 + B_{k0}^{*2} = S_1^{(z+1)2}$ ist. Dann ist die F-Zahl von $S_2^{(z+1)2} + \cdots + S_{m^*}^{(z+1)2}$ gleich der von $S_2^{(z)2} + \cdots + S_{m^*}^{(z)2}$, und $S_1^{(z+1)2}$ ist bezüglich Θ ähnlich zu der 2-Sphäre $B_{k0}^2 + X_k^2 + B_{kt_k+1}^2$. Diese 2-Sphäre ist aber eine z-Komponente der 2-Mannigfaltigkeit $\sigma_s \ldots \sigma_{k+1} \bar{\sigma}_k \sigma_{k-1} \ldots \sigma_1 (M_1^2 + M_2^2)$, wenn $\bar{\sigma}_k$ die zu σ_k entgegengesetzte Umschaltung bedeutet. Diese 2-Mannigfaltigkeit besitzt nach [5], Hauptsatz 4 dieselbe F-Zahl wie

S_1^{*2} und enthält eine Falte. Also ist entweder $f^{(z+1)} < f^{(z)}$, oder es ist $f^{(z+1)} = f^{(z)}$ und $S_1^{(z+1)2}$ enthält eine Falte, womit das Sphärensystem die geforderten Eigenschaften (für das Vorliegen von Fall I) besitzt.

Mit 3.5.1, 3.5.2 und 3.5.3 ist die Existenz des Sphärensystems $S_1^{(z+1)2} + \cdots + S_{m^*}^{(z+1)2}$ für die unter 3.3 genannten Fälle nachgewiesen, deren einer stets vorliegt. Damit ist der Beweis des Hilfssatzes 2 vollendet.

BESTIMMUNGSVERFAHREN 2. *Ist M^3 eine kreuzhauben-, henkel- und randsphärenfreie, sonst beliebige (zusammenhängende, berandete oder unberandete) 3-Mannigfaltigkeit, so läßt sich (in einer beschränkten Anzahl von Schritten) ein vollständiges Sphärensystem in M^3 konstruieren und damit eine Aufspaltung der 3-Mannigfaltigkeit M^3 bzw. ihrer Äquivalenzklasse $\mathfrak{M}^3$ in irreduzible Erzeugende angeben.*

Dieses kann folgendermaßen geschehen: Es wird eine Normalzerlegung Θ bezüglich einer beliebigen randtreuen Zerlegung Γ von M^3 in offene Raumelemente gewählt. Dann wird aus jeder der Fundamentalklassen bezüglich Θ ein Repräsentant konstruiert, und unter diesen Repräsentanten werden diejenigen ausgewählt, die 2-Sphären sind; ihre P-Zahlenvektoren seien $\mathfrak{x}_1, \ldots, \mathfrak{x}_t$. Diese 2-Sphären werden ohne Veränderung ihrer Ähnlichkeitsklassen so deformiert, daß je zwei von ihnen entweder miteinander punktfremd werden, oder einander nur in geschlossenen Linien durchdringen. Die so erhaltenen 2-Sphären seien $S_1^2, \ldots, S_t^2$. Hierbei läßt sich erreichen, daß folgendes gilt: Gibt es zu $S_{i_1}^2$, $S_{i_2}^2$ (i_1, i_2 zwischen 1 und t) zwei bezüglich Θ ähnliche 2-Sphären, die miteinander punktfremd sind, so sind auch bereits $S_{i_1}^2$, $S_{i_2}^2$ miteinander punktfremd. Enthält eine 2-Sphäre S_i^2 Durchdringungslinien mit anderen 2-Sphären S_i^2, so beseitigen wir diese Durchdringungslinien durch ϱ-Operationen und gelangen nach einer beschränkten Anzahl von Schritten zu einem System von miteinander punktfremden 2-Sphären $S_1'^2, \ldots, S_u'^2$. Dabei lassen sich die ϱ-Operationen so wählen, daß folgendes gilt: Sind $S_1^2, \ldots, S_t^2$ durch die ersten k ϱ-Operationen in $S_{(k)1}^2, \ldots, S_{(k)t^*}^2$ übergegangen und wird $S_{(k)1}^2$ durch die $(k+1)$-te ϱ-Operation abgeändert, so ist der Kern dieser ϱ-Operation mit allen denjenigen $S_{(k)i}^2$ (i zwischen 1 und t^*) punktfremd, mit denen $S_{(k)1}^2$ punktfremd ist. Unter den 2-Sphären können wir diejenigen fortlassen, die sich als kompressibel oder als geometrisch von den übrigen abhängig erkennen lassen. Können alle 2-Sphären in dieser Weise fortgelassen werden, so ist M^3 irreduzibel; anderenfalls bilden die restlichen 2-Sphären $S_1''^2, \ldots, S_v''^2$ ein vollständiges Sphärensystem in M^3 und die Ergänzung der z-Komponenten von $M^3 - (S_1''^2 + \cdots + S_v''^2)$ durch 3-Elemente $E_1^3, E_1^{*3}, \ldots, E_v^3, E_v^{*3}$ mit $E_i^3 = E_i^{*3} = S_i''^2$ (für alle $i = 1, \ldots, v$), deren Innenteile miteinander und mit M^3 punktfremd sind, ergibt eine Aufspaltung von M^3 in irreduzible Erzeugende.

BEWEIS. Es ist noch zu beweisen, daß das Sphärensystem $S_1''^2 + \cdots + S_v'^{2'}$ in M^3 vollständig ist: Aus dem Endlichkeitssatz 1 folgt, daß ein vollständiges Sphärensystem in M^3 existiert. Daraus folgt nach [5], Hauptsatz 1 und nach Satz 3, daß ein vollständiges Sphärensystem $S_1^{(1)2} + \cdots + S_m^{(1)2}$ in M^3

existiert, das eine Normalfläche bezüglich Θ ist; ihre F-Zahl sei $f^{(1)}$. Sind nun die P-Zahlenvektoren von $S_1^{(1)2}, \ldots, S_m^{(1)2}$ nicht sämtlich Fundamentallösungen der P-Gleichungen, so existiert nach Hilfssatz 2 ein vollständiges Sphärensystem $S_1^{(2)2}+\cdots+S_{m^{(2)}}^{(2)2}$ in M^3, das ebenfalls eine Normalfläche ist, und dessen F-Zahl $f^{(2)}$ kleiner ist als $f^{(1)}$. So läßt sich weiter auf die Existenz eines vollständigen Sphärensystems $S_1^{*2}, \ldots, S_{m^*}^{*2}$ in M^3 schließen, so daß jede der 2-Sphären in einer der Fundamentalklassen liegt und keine zwei von ihnen in derselben Klasse liegen (denn ähnliche 2-Sphären sind geometrisch abhängig). Unter den oben konstruierten 2-Sphären $S_1^2, \ldots, S_t^2$ befinden sich also m^* — sie mögen mit $S_{i_1}^2, \ldots, S_{i_{m^*}}^2$ bezeichnet werden —, deren Summe zu $S_1^{*2}+\cdots+S_{m^*}^{*2}$ bezüglich Θ ähnlich ist, die also ein vollständiges Sphärensystem in M^3 bilden. Bei den Operationen, durch die die 2-Sphären $S_1^2, \ldots, S_t^2$ in die miteinander punktfremden 2-Sphären $S_1'^2, \ldots, S_u'^2$ übergeführt werden, wird das Sphärensystem $S_{i_1}^2+\cdots+S_{i_{m^*}}^2$, wenn überhaupt, nur durch ϱ-Operationen verändert, geht dabei also nach Satz 3 in ein vollständiges Sphärensystem über. Da dieses Sphärensystem in $S_1'^2+\cdots+S_u'^2$ liegt, ist also auch $S_1'^2+\cdots+S_u'^2$ vollständig, und damit auch das durch Fortlassen geometrisch abhängiger 2-Sphären entstehende Sphärensystem $S_1''^2+\cdots+S_v''^2$, w.z.b.w.

Folgerung 1. *Ist eine beliebige (zusammenhängende, berandete oder unberandete) 3-Mannigfaltigkeit M^3 gegeben, so läßt sich (in einer beschränkten Anzahl von Schritten) ein vollständiges Sphärensystem in M^3 konstruieren und damit eine Aufspaltung von M^3 bzw. $\mathfrak{M}^3$ in irreduzible Erzeugende angeben. Ferner lassen sich Henkel- und Kreuzhaubenzahl von M^3 bestimmen.*

Hierzu wird zunächst ebenso vorgegangen, wie beim Bestimmungsverfahren 1: Durch Reduktion der Kreuzhauben $K_1^3, \ldots, K_k^3$ und der Henkel $H_1^3, \ldots, H_h^3$ und Ausfüllen der Randsphären wird eine kreuzhauben-, henkel- und randsphärenfreie 3-Mannigfaltigkeit M'^3 hergestellt. (k und h sind Kreuzhauben- und Henkelzahl von M^3.) In M'^3 läßt sich nach Bestimmungsverfahren 2 ein vollständiges Sphärensystem $S_1^2+\cdots+S_v^2$ konstruieren. Dabei läßt sich durch isotope Deformationen erreichen, daß dieses Sphärensystem in $M^3-(K_1^3+\cdots+K_k^3+H_1^3+\cdots+H_h^3)$ liegt. Dann ist $S_1^2+\cdots+S_v^2+\dot{K}_1^3+\cdots+\dot{K}_k^3+\dot{H}_1^3+\cdots+\dot{H}_h^3$ ein vollständiges Sphärensystem in M^3 (das aus $v+k+2h$ 2-Sphären besteht); denn jede mit diesen 2-Sphären punktfremde 2-Sphäre in M^3 liegt entweder in $M^3-(K_1^3+\cdots+K_k^3+H_1^3+\cdots+H_h^3)$ (und ist von $S_1^2, \ldots, \dot{H}_h^3$ geometrisch abhängig) oder in einer der Kreuzhauben K_i^3 (und ist von $\dot{K}_i^3$ geometrisch abhängig, da der projektive Raum nach Satz 7 irreduzibel ist) oder in einer der zweifach gelochten 3-Sphären H_i^3 (und ist von $\dot{H}_i^3$ geometrisch abhängig).

Folgerung 2. *Durch die Bestimmungsverfahren 1 und 2 und die Eindeutigkeitssätze 1 und 2 ist das Homöomorphieproblem der (kompakten, berandeten oder unberandeten) 3-Mannigfaltigkeiten auf das der irreduziblen 3-Mannigfaltigkeiten zurückgeführt (im Falle orientierbarer 3-Mannigfaltigkeiten auf das der orientierten irreduziblen 3-Mannigfaltigkeiten).*

Denn sind zwei kompakte 3-Mannigfaltigkeiten M_1^3 und M_2^3 vorgelegt, so lassen sich nach Folgerung 1 die Henkelzahlen, je eine Aufspaltung in irreduzible Erzeugende (und trivialer Weise auch die Randsphärenzahlen und die Orientierbarkeitscharaktere) von M_1^3 und M_2^3 bestimmen. Damit läßt sich feststellen, ob die Henkelzahlen, Randsphärenzahlen und Orientierbarkeitscharaktere übereinstimmen; die beiden 3-Mannigfaltigkeiten sind nach den Eindeutigkeitssätzen 1 und 2 dann und nur dann homöomorph, wenn dieses der Fall ist und wenn unter den konstruierten irreduziblen Erzeugenden diejenigen paarweise homöomorph sind (bei orientierbaren M_1^3, M_2^3 mit entsprechenden Orientierungen), die keine 3-Sphären sind.

FOLGERUNG 3. *Ist M^3 eine beliebige (zusammenhängende, berandete oder unberandete) 3-Mannigfaltigkeit, und sind in M^3 beliebige, miteinander punktfremde, zusammenhängende abgeschlossene Punktmengen $P_1, \ldots, P_s$ gegeben, so läßt sich (in einer beschränkten Anzahl von Schritten) entscheiden, ob $P_1, \ldots, P_s$ in M^3 miteinander verkettet seien. Insbesondere läßt sich in M^3 ein mit $P_1, \ldots, P_s$ punktfremdes Sphärensystem $S_1^2 + \cdots + S_w^2$ konstruieren, so daß folgendes gilt: Die Punktmengen $P_{i_1}, \ldots, P_{i_t}$ ($i_1, \ldots, i_t$ beliebig zwischen 1 und s) sind dann und nur dann miteinander verkettet in M^3, wenn sie in einer und derselben z-Komponente von $M^3 - (S_1^2 + \cdots + S_w^2)$ liegen.*

Hierzu werden in M^3 kleine Umgebungen $U_1^3, \ldots, U_s^3$ von $P_1, \ldots, P_s$ gewählt. Die 3-Mannigfaltigkeit $M^3 - (\mathring{U}_1^3 + \cdots + \mathring{U}_s^3)$ ist nicht notwendig zusammenhängend; ihre z-Komponenten bezeichnen wir mit $M_1^3, \ldots, M_p^3$. In den 3-Mannigfaltigkeiten $M_1^3, \ldots, M_p^3$ konstruieren wir, wie in Folgerung 1 beschrieben, je ein vollständiges Sphärensystem $S_{11}^2 + \cdots + S_{1v_1}^2, \ldots, S_{p1}^2 + \cdots + S_{pv_p}^2$. Liegen in M_j^3 ($j = 1, \ldots, p$) 2-Sphären, so werden diese mit $S_{j1}'^2, \ldots, S_{jg_j}'^2$ bezeichnet, und zu jeder von ihnen wird in M_j^3 eine nahe benachbarte 2-Sphäre $S_{j1}^{*2}, \ldots, S_{jg_j}^{*2}$ gewählt. Dann ist $S_{11}^2 + \cdots + S_{pv_p}^2 + S_{11}^{*2} + \cdots + S_{pg_p}^{*2}$ ein Sphärensystem — bei einfacher Numerierung schreiben wir es als $S_1^2 + \cdots + S_w^2$ — mit den geforderten Eigenschaften.

BEWEIS. Es genügt, folgendes zu beweisen: Fügen wir in M^3 zu dem Sphärensystem $S_1^2 + \cdots + S_w^2$ eine beliebige Anzahl von 2-Sphären $S_{w+1}^2, \ldots, S_{w+q}^2$ hinzu, die miteinander, mit $S_1^2, \ldots, S_w^2$ und mit $U_1^3, \ldots, U_s^3$ punktfremd sind, so gilt: liegen zwei Punktmengen P_{i_1}, P_{i_2} (i_1, i_2 beliebig zwischen 1 und s) in einer und derselben z-Komponente von $M^3 - (S_1^2 + \cdots + S_w^2)$, so liegen sie auch in einer und derselben z-Komponente von $M^3 - (S_1^2 + \cdots + S_{w+q}^2)$. — Wir machen die Induktionsannahme, die Behauptung sei für $q = r$ bewiesen ($r = 0, 1, 2, \ldots$); dann gilt für $q = r+1$: S_{w+r+1}^2 liegt in einer der 3-Mannigfaltigkeiten $M_1^3, \ldots, M_p^3$ — etwa in M_l^3 —, also in einer z-Komponente Z^3 von $M_l^3 - (S_1^2 + \cdots + S_{w+r}^2)$. Da der in M_l^3 liegende Teil des Sphärensystems $S_1^2 + \cdots + S_{w+r}^2$ ein vollständiges Sphärensystem in M_l^3 ist, wird Z^3 durch S_{w+r+1}^2 in zwei z-Komponenten Z'^3 und Z''^3 zerlegt, deren mindestens eine — diese sei Z'^3 — eine gelochte 3-Sphäre ist. Nun ist Z'^3 entweder mit allen Umgebungen $U_1^3, \ldots, U_s^3$ punktfremd, oder Z'^3 enthält genau eine der 2-Sphären $S_{lm}'^2$ (m zwischen 1 und g_l). Im letzteren Falle ist Z''^3 mit den Umgebungen

$U_1^3, \ldots, U_s^3$ punktfremd; denn dann kann Z^3 außer $S_{l_m}'^2$ keine weiteren Punkte aus $U_1^3, \ldots, U_s^3$ enthalten (weil sich unter den $S_1^2, \ldots, S_w^2$ die 2-Sphäre $S_{l_m}^{*2}$ befindet, die in M_l^3 zusammen mit $S_{l_m}'^2$ eine zweifach gelochte 3-Sphäre berandet, in der Z^3 liegt). Also ist in jedem Falle mindestens eine der z-Komponenten Z'^3, Z''^3 — sie werde mit Z^{*3} bezeichnet — mit $U_1^3, \ldots, U_s^3$ punktfremd. — Nun werde die z-Komponente von $M^3-(S_1^2+\cdots+S_{w+r}^2)$, in der S_{w+r+1}^2 liegt, mit T^3 bezeichnet. Dann wird T^3 durch S_{w+r+1}^2 in zwei z-Komponenten zerlegt, deren eine Z^{*3} ist; denn da Z^{*3} eine mit $U_1^3, \ldots, U_s^3$ punktfremde z-Komponente von $M_l^3-(S_1^2+\cdots+S_{w+r+1}^2)$ ist, ist Z^{*3} auch eine z-Komponente von $M^3-(S_1^2+\cdots+S_{w+r+1}^2)$. Die andere z-Komponente von $T^3-S_{w+r+1}^2$ werde mit T'^3 bezeichnet. Liegen nun P_{i_1} und P_{i_2} in einer von T^3 verschiedenen z-Komponente von $M^3-(S_1^2+\cdots+S_{w+r}^2)$, so ist diese auch eine z-Komponente von $M^3-(S_1^2+\cdots+S_{w+r+1}^2)$; liegen P_{i_1} und P_{i_2} in T^3, so liegen sie in T'^3, (da Z^{*3} mit $P_1, \ldots, P_s$ punktfremd ist). Damit ist die Behauptung für $q=r+1$ bewiesen; da sie für $q=0$ trivial ist, ist der Induktionsbeweis damit vollendet.

Literatur

[1] Alexander, J. W.: On the subdivision of polyhedra in 3-space. Proc. Nat. Acad. Sci. **10**, 6–8 (1924).

[2] Alexandroff, P., u. H. Hopf: Topologie. Berlin 1935.

[3] Bing, R. H.: An alternative proof that 3-manifolds can be triangulated. Ann. Math. **69**, 37–65 (1959).

[4] Fox, R. H.: On the imbedding of polyhedra in 3-space. Ann. Math. **49**, 462–470 (1948).

[5] Haken, W.: Theorie der Normalflächen. Ein Isotopiekriterium für den Kreisknoten. Acta math. **105**, 245–375 (1961).

[6] Kneser, H.: Geschlossene Flächen in dreidimensionalen Mannigfaltigkeiten. Jber. dtsch. Mathematikervereinig. **38**, 248–260 (1929).

[7] Milnor, J. W.: A unique decomposition theorem for 3-manifolds. Im Druck.

[8] Moise, E. E.: Affine structures in 3-manifolds. V. Ann. Math. **56**, 96–114 (1952).

[9] — Affine Structures in 3-manifolds. VIII. Ann. Math. **59**, 159–170 (1954).

[10] Papakyriakopoulos, C. D.: On Dehn's lemma and the asphericity of knots. Ann. Math. **66**, 1–26 (1957).

[11] — Some problems on 3-dimensional manifolds. Bull. Amer. Math. Soc. **64**, 317–335 (1958).

[12] Reidemeister, K.: Topologie der Polyeder und kombinatorische Topologie der Komplexe. Leipzig 1953.

[13] Seifert, H., u. W. Threlfall: Lehrbuch der Topologie. Leipzig 1934.

[14] Whitehead, J. H. C.: On 2-spheres in 3-manifolds. Bull. Amer. Math. Soc. **64**, 161–166 (1958).

München, Winthirplatz 5

(Eingegangen am 15. Dezember 1960)

BIERLEIN, D.
Math. Zeitschr. 76, 468–471 (1961)

Der Graph meßbarer Funktionen mit abstraktem Definitionsbereich

Von
DIETER BIERLEIN

Für eine reelle Funktion f, die auf einem *topologischen* Raum M definiert ist, sind Beziehungen zwischen der Meßbarkeit der Funktion und ihrem Graphen $X(f) := \{(x, f(x)) : x \in M\}$ bekannt oder unschwer aus bekannten Sätzen zu gewinnen. Hinreichend für die Meßbarkeit von f ist die Bedingung

(*) $X(f)$ ist eine Suslinsche analytische Menge bezüglich der Produkttopologie auf der Produktmenge $(M, \boldsymbol{R}^1)$,

falls M noch bestimmten — von der Art der Meßbarkeit abhängigen — topologischen Voraussetzungen unterworfen ist[1]). So ist z. B. unter der recht kräftigen Einschränkung, daß M eine separable, absolut analytische Menge ist, die Bedingung (*) hinreichend dafür, daß f Borel-meßbar (Bairesch) ist[2]). Die Theorie der Kapazitäten über topologischen Räumen[3]) bietet die Möglichkeit, schwächere Bedingungen für M zu finden, die zusammen mit (*) hinreichend sind für die Meßbarkeit von f bezüglich der Vervollständigung eines auf dem σ-Körper der Borelschen Mengen definierten Maßes; dabei ist die Tatsache auszunützen, daß das äußere Maß eine Kapazität ist.

Ein neueres Ergebnis von CHOQUET über abstrakte Kapazitäten[4]) erlaubt es — wie in dieser Arbeit ausgeführt wird — sich von topologischen Voraussetzungen völlig freizumachen und im Fall eines beliebigen vollständigen Maßfeldes $(M, \mathfrak{R}, \mu)$ eine zu (*) analoge Bedingung als hinreichend für die μ-Meßbarkeit von f zu gewinnen, wodurch sich zugleich erweist, daß die Voraussetzungen über M, die für die Anwendung der („klassischen") Theorie der Kapazitäten über topologischen Räumen zu beachten sind, tatsächlich überflüssig sind. Unter einem vollständigen Maßfeld soll dabei ein Tripel aus einer Menge M, einem σ-Körper $\mathfrak{R}$ von Teilmengen von M und einem vollständigen (σ-additiven) Maß[5]) μ mit $\mathfrak{R}$ als Definitionsbereich verstanden werden.

Für den Zusammenhang zwischen Graph und Meßbarkeit einer Funktion ergibt sich aus Satz 1 dann folgendes Bild:

$f | M$ ist genau dann μ-meßbar, wenn $X(f)$ ein $\overline{R}_\sigma$ ist, d.h. wenn der Graph sich als Komplement einer abzählbaren Vereinigung von „Rechtecken"

[1]) Vgl. [2], [3], [5] u.a.
[2]) Vgl. [5], Satz 43.2.1.
[3]) Vgl. [3].
[4]) Vgl. [4].
[5]) Im Sinn von [6], (I 3.4).

$R=(K, B)$ mit „Seiten" K aus $\mathfrak{R}$ und B aus $\mathfrak{B}$, dem σ-Körper der Borel-Mengen des $\boldsymbol{R}^1$, darstellen läßt. Dabei stimmt die Menge der Graphen, die sich als Kern eines Suslin-Schemas[6]) aus Rechtecken der Form (K, B) darstellen lassen (kurz: $\mathfrak{R}$-Suslinsche Graphen), mit der Menge der Graphen $X(f)\in{}^{B}(\mathfrak{R}\times\mathfrak{B})$ und mit der Menge der Graphen vom Typ $\overline{R_\sigma}$ überein. Es gibt also überhaupt keine Funktion, deren Graph $\mathfrak{R}$-Suslinsch ist oder zu ${}^{B}(\mathfrak{R}\times\mathfrak{B})$ gehört, aber nicht vom Typ $\overline{R_\sigma}$ ist. Anders formuliert:

Ist f μ-meßbar, so ist $X(f)$ sogar ein $\overline{R_\sigma}$,

ist f nicht μ-meßbar, so ist $X(f)$ nicht einmal $\mathfrak{R}$-Suslinsch.

Bezeichnungen und Definitionen:

$\boldsymbol{N}$ ist die Menge der natürlichen Zahlen,
$\boldsymbol{F}:=\boldsymbol{N}^{\boldsymbol{N}}$ die Menge der Folgen natürlicher Zahlen;
für $\boldsymbol{F}\ni\sigma=(n_1, n_2, \ldots)$ setzen wir $n_k(\sigma):=(n_1, n_2, \ldots, n_k)$.

Ist $\mathfrak{L}$ ein System von Teilmengen einer festen Menge, dann sei

${}^{R}\mathfrak{L}$ der kleinste $\sigma\delta$-Ring[7]), der $\mathfrak{L}$ umfaßt
(in [1]: ${}^{R}\mathfrak{L}=\mathfrak{L}^{B}$, „kleinstes Borelsches System über $\mathfrak{L}$"),

${}^{B}\mathfrak{L}$ der kleinste σ-Körper, der $\mathfrak{L}$ umfaßt
(„Borelsche Erweiterung von $\mathfrak{L}$"),

${}^{S}\mathfrak{L}$ das System der Suslin-Kerne zu $\mathfrak{L}$
(„Suslinsche Erweiterung von $\mathfrak{L}$"; in [4]: System der „$\mathfrak{L}$-Suslinschen" Mengen).

Wie üblich ist dabei definiert: Ein *Suslin-Schema* zu $\mathfrak{L}$ ist ein System $(L_n)_{n\in\boldsymbol{N}^k,\,k\in\boldsymbol{N}}$ mit $L_n\in\mathfrak{L}$; der *Kern* des Schemas (L_n) ist die Menge $\sum_{\sigma\in\boldsymbol{F}}\prod_{k\in\boldsymbol{N}} L_{n_k(\sigma)}$.

Es sei daran erinnert, daß $\mathfrak{L}\subset{}^{R}\mathfrak{L}\subset{}^{S}\mathfrak{L}={}^{S}({}^{S}\mathfrak{L})$ allgemein gilt und ${}^{R}\mathfrak{L}={}^{B}\mathfrak{L}$ dann, wenn $\overline{L}\in{}^{R}\mathfrak{L}$ für jedes $L\in\mathfrak{L}$[8]).

Ist $A\subset(M_1, M_2)$, dann bezeichne $pr_{M_k}A$ die Projektion von A auf M_k. Ist $\mathfrak{L}_\varkappa$ ein System von Teilmengen von $M_\varkappa$, dann sei

$$\mathfrak{R}(\mathfrak{L}_1, \mathfrak{L}_2):=\{(L_1, L_2): L_\varkappa\in\mathfrak{L}_\varkappa\},$$

also das System der „Rechtecke" mit „Seiten" aus $\mathfrak{L}_1$ bzw. $\mathfrak{L}_2$.

Zum Beweis des angekündigten Satzes verwenden wir zwei Hilfssätze:

Hilfssatz 1. *Ist $(M, \mathfrak{R}, \mu)$ ein vollständiges Maßfeld, so ist ${}^{S}\mathfrak{R}=\mathfrak{R}$.*

Beweis. Das zu μ gehörige äußere Maß μ^* ist eine abstrakte Kapazität über $(M, \mathfrak{R})$ im Sinn von Choquet [4], S. 84. Eine Menge $A\subset M$ ist in diesem Sinne genau dann $(\mu^*, \mathfrak{R})$-kapazitabel, wenn $\mu_*(A)=\mu^*(A)$ ist; wegen der Vollständigkeit von μ also genau dann, wenn A μ-meßbar ist; d.h. im Fall $A\in\mathfrak{R}$. Da alle Mengen aus ${}^{S}\mathfrak{R}$ nach [4], Théorème 1, $(\mu^*, \mathfrak{R})$-kapazitabel sind, folgt ${}^{S}\mathfrak{R}\subset\mathfrak{R}$ und damit wegen der allgemeinen Beziehung $\mathfrak{L}\subset{}^{S}\mathfrak{L}$ die Behauptung.

[6]) Definition s. unten.
[7]) $\mathfrak{L}_1$ heißt $\sigma\delta$-Ring, falls mit $L_n\in\mathfrak{L}_1$ $(n\in\boldsymbol{N})$ stets auch $\sum_n L_n$ und $\prod_n L_n\in\mathfrak{L}_1$.
[8]) Vgl. z.B. [1], 2.9.

HILFSSATZ 2. *$\mathfrak{R}$ sei ein σ-Körper von Teilmengen von M; dann gelten die Beziehungen*

a) ${}^{B}(\mathfrak{R}\times\mathfrak{B})\subset{}^{S}\mathscr{R}(\mathfrak{R},\mathfrak{B})$,

b) *Für* $A\in{}^{S}\mathscr{R}(\mathfrak{R},\mathfrak{B})$ *folgt* $pr_M A\in{}^{S}\mathfrak{R}$.

BEWEIS. 1. Es sei $\mathfrak{J}:=\{[a,b]\colon -\infty<a\leqq b<\infty\}$; dann gilt wegen ${}^{B}\mathfrak{J}=\mathfrak{B}$ zunächst $\mathscr{R}(\mathfrak{R},\mathfrak{B})\subset{}^{B}\mathscr{R}(\mathfrak{R},\mathfrak{J})$. Da für $R=(K,J)\in\mathscr{R}(\mathfrak{R},\mathfrak{J})$ stets $\overline{R}=(\overline{K},J)+(M,\overline{J})\in{}^{R}\mathscr{R}(\mathfrak{R},\mathfrak{J})$, gilt ${}^{R}\mathscr{R}(\mathfrak{R},\mathfrak{J})={}^{B}\mathscr{R}(\mathfrak{R},\mathfrak{J})$. Daraus erhalten wir

$$\mathscr{R}(\mathfrak{R},\mathfrak{B})\subset{}^{B}\mathscr{R}(\mathfrak{R},\mathfrak{J})={}^{R}\mathscr{R}(\mathfrak{R},\mathfrak{J})\subset{}^{S}\mathscr{R}(\mathfrak{R},\mathfrak{J})\subset{}^{S}\mathscr{R}(\mathfrak{R},\mathfrak{B}). \tag{1}$$

Gehen wir in (1) gliedweise zur Suslinschen Erweiterung über, so gelangen wir wegen der allgemeinen Relation ${}^{S}({}^{S}\mathfrak{Q})={}^{S}\mathfrak{Q}$ zu

$${}^{S}\mathscr{R}(\mathfrak{R},\mathfrak{B})={}^{S}\mathscr{R}(\mathfrak{R},\mathfrak{J}). \tag{2}$$

Aus (1) erhält man wegen ${}^{B}\mathscr{R}(\mathfrak{R},\mathfrak{J})={}^{B}(\mathfrak{R}\times\mathfrak{B})$ unmittelbar Beh. a).

2. Zu $A\in{}^{S}\mathscr{R}(\mathfrak{R},\mathfrak{B})$ existiert wegen (2) ein Suslin-Schema $(R_{\mathfrak{n}})_{\mathfrak{n}\in\mathbf{N}^{k},\,k\in\mathbf{N}}$ mit den Eigenschaften

α) $R_{\mathfrak{n}}=(K_{\mathfrak{n}},J_{\mathfrak{n}})\in\mathscr{R}(\mathfrak{R},\mathfrak{J})$,

β) $A=\sum\limits_{\sigma\in\boldsymbol{F}}^{\cdot}\prod\limits_{k\in\boldsymbol{N}}^{\cdot}R_{\mathfrak{n}_k(\sigma)}$,

wobei wir o.E.d.A. folgende Normierungen vornehmen:

γ) $R_{\mathfrak{n}_1,\ldots,\mathfrak{n}_{k+1}}\subset R_{\mathfrak{n}_1,\ldots,\mathfrak{n}_k}$,

δ) $K_{\mathfrak{n}}\dot{-}J_{\mathfrak{n}}=0$ für $R_{\mathfrak{n}}=0$.

Wegen (γ) und (δ) folgt

$$J_{\mathfrak{n}_1,\ldots,\mathfrak{n}_{k+1}}\subset J_{\mathfrak{n}_1,\ldots,\mathfrak{n}_k}. \tag{3}$$

Für jedes $\sigma\in\boldsymbol{F}$ sei $R_\sigma:=\prod\limits_{k}^{\cdot}R_{\mathfrak{n}_k(\sigma)}$; analog definieren wir K_σ und J_σ; endlich sei $\boldsymbol{F}_1:=\{\sigma\colon J_\sigma\neq 0\}$.

Dann ist $A=\sum\limits_{\sigma\in\boldsymbol{F}}^{\cdot}R_\sigma$, $R_\sigma=(K_\sigma,J_\sigma)$ und folglich

$$pr_M A=\sum_{\sigma\in\boldsymbol{F}}^{\cdot}pr_M R_\sigma=\sum_{\sigma\in\boldsymbol{F}_1}^{\cdot}K_\sigma. \tag{4}$$

Zu jedem $\sigma\in\boldsymbol{F}-\boldsymbol{F}_1$ existiert wegen (3) ein $k_0=k_0(\sigma)$, so daß $J_{\mathfrak{n}_0}=0$ für $\mathfrak{n}_0:=\mathfrak{n}_{k_0}(\sigma)$; daraus folgt $R_{\mathfrak{n}_0}=0$, wegen (δ) dann auch $K_{\mathfrak{n}_0}=0$ und deshalb $K_\sigma=0$.

Somit geht (4) über in $pr_M A=\sum\limits_{\sigma\in\boldsymbol{F}}^{\cdot}\prod\limits_{k\in\boldsymbol{N}}^{\cdot}K_{\mathfrak{n}_k(\sigma)}$. Dies bedeutet gerade, daß $pr_M A$ Kern des Suslin-Schemas $(K_{\mathfrak{n}})$ ist, also zu ${}^{S}\mathfrak{R}$ gehört: Beh. b).

SATZ 1. *Ist $(M,\mathfrak{R},\mu)$ ein vollständiges Maßfeld und $\mathfrak{R}=\mathscr{R}(\mathfrak{R},\mathfrak{B})$, dann sind die folgenden vier Aussagen äquivalent:*

(1) *f ist μ-meßbar ($\mathfrak{R}$-meßbar)*

(2) *$X(f)$ ist ein $\overline{R}_\sigma$, d.h. es existieren abzählbar viele „Rechtecke" $R_{\mathfrak{n}}\in\mathfrak{R}$, so daß $\overline{X(f)}=\sum\limits_{\mathfrak{n}}^{\cdot}R_{\mathfrak{n}}$*

(3) $X(f) \in {}^{B}(\mathfrak{R} \times \mathfrak{B})$

(4) $X(f)$ *ist* $\mathfrak{R}$*-Suslinsch, d.h.* $X[f] \in {}^{S}\mathfrak{R}$.

BEWEIS. Gilt (1), so folgt $\{x: f(x) \gtrless y\} \in \mathfrak{R}$ für jedes $y \in \boldsymbol{R}^1$ und

$$\overline{X(f)} = \{(x, y): f(x) > y\} + \{(x, y): f(x) < y\}$$
$$= \sum_{r \text{ rat.}}{}^{*} (\{x: f(x) > r\},\ (-\infty, r]) + \cdots;$$

$\overline{X(f)}$ ist also ein R_σ: Aussage (2).

Aus (2) folgt (3) unmittelbar und aus (3) nach Hilfssatz 2, a) auch (4).

Gilt (4), so ist $X(f) \cdot \{(x, y): y \leqq y_0\} \in {}^{S}\mathfrak{R}$ für jedes $y_0 \in \boldsymbol{R}^1$ und deshalb wegen Hilfssatz 2, b)

$$\{x: f(x) \leqq y_0\} = pr_M \left(X(f) \cdot \{(x, y): y \leqq y_0\}\right) \in {}^{S}\mathfrak{R};$$

daraus folgt mit Hilfssatz 1 wieder (1).

Literatur

[1] AUMANN, G.: Reelle Funktionen. Berlin 1954.

[2] BOURBAKI, N.: Topologie Générale, Kap. 9. Paris 1958.

[3] CHOQUET, G.: Theory of capacities. Ann. Inst. Fourier **5** (1953).

[4] — Forme abstraite du théorème de capacitabilité. Ann. Inst. Fourier **9** (1959).

[5] HAHN, H.: Reelle Funktionen, Teil 1. Leipzig 1932.

[6] RICHTER, H.: Wahrscheinlichkeitstheorie. Göttingen 1956.

München, Mathematisches Institut der Universität

(Eingegangen am 25. März 1961)